

Compound and Other Optimum Designs for Systems of Nonlinear Differential Equations Arising in Chemical Kinetics

Anthony C. Atkinson, Department of Statistics,
London School of Economics, London WC2A 2AE, UK
and

Barbara Bogacka, School of Mathematical Sciences,
Queen Mary and Westfield College, London E1 4NS

August 26, 2001

Abstract

The optimum design of experiments for nonlinear models requires parameter sensitivities, that is the derivatives of the response with respect to the parameters. If the differential equations forming the kinetic model do not have an analytical solution, numerical derivatives have to be used. We describe the “direct” method for calculating the sensitivities and apply it to the design of experiments for estimating the order of chemical reactions.

Keywords: Bayesian design; multipurpose design; parameter estimation; sensitivity.

1 Introduction

Kinetic models of chemical reactions are in the form of differential equations containing two groups of parameters, the rates and orders of the reactions. Nearly all work in the statistical literature on the design of experiments for estimation of these parameters assumes that the orders of the reactions are known and that the rates are to be estimated. In simple cases the differential equations can then be solved to give analytical expressions for the expected response to be measured in the experiment. The resulting relationship between response and parameters is nonlinear and there is a large literature

on the design of experiments for such nonlinear models. An introduction is given in Chapter 18 of Atkinson and Donev (1992). The emphasis in the present paper, on the other hand, is on the more complicated cases when the differential equations cannot be solved analytically.

The statistical literature on designs for nonlinear models starts with Box and Lucas (1959) who find D-optimum designs for the rates of reaction in several models including first-order decay and two consecutive first-order reactions. The calculation of these designs requires the derivatives of the expected response with respect to the parameters, which are found by analytical differentiation. These derivatives are also required by numerically efficient methods for estimation of the kinetic parameters such as the Gauss-Newton least squares algorithm. If the differential equations of the kinetic model have to be solved numerically, numerical methods also have to be used to find the derivatives. There is an appreciable chemical literature on the solution of these differential equations, including methods of finding the derivatives, which is focused on the analysis of data, not on the design of experiments.

In order to find the required derivatives we use what Valko and Vajda (1984) call the “direct” method which leads to additional differential equations for the required derivatives. These new equations have also to be solved numerically. A statistical reference is p.178 of Bates and Watts (1988), who are concerned with data analysis, not design. There are however a few examples of the use of the direct method to find D-optimum designs for experiments in biomedical and chemical engineering. We describe these experiments in Section 8.

We consider the design of experiments when the order of the chemical reactions is to be estimated as well as the rate, an important problem which has received little treatment in the statistical literature. For all but the simplest models we obtain systems of equations which have to be solved numerically. The simplest case of all is that of the extension of the first-order decay model considered by Box and Lucas to general decay, in which the order of the reaction is a parameter of interest. Atkinson and Bogacka (1997) find a variety of designs, including compound designs, mainly for the determination of the order. In this case it is again possible to find an analytical expression for the response and so for the required derivatives. However the extension of the model for two consecutive reactions to general order leads to a set of differential equations which cannot be solved analytically. The required derivatives then have to be found numerically. The present paper is concerned with the design of experiments for this model when both the reaction rate and the order matter.

Similar calculation of numerical derivatives is required for D-optimum

designs for reversible reactions, such as those described by Atkinson, Bogacka, and Bogacki (1998). They confine themselves to first-order kinetics, as do Xu, Liang, and Fang (2000) who compare D-optimum designs with those provided by other criteria in the presence of parameter uncertainty. The extension to other orders of reaction is not considered by either set of authors. Such extension, and the calculation of the T-optimum designs for model discrimination found by Atkinson et al. (1998), are greatly simplified by the use of the direct method. As an illustration we give the differential equations for the derivatives for one reversible reaction in Appendix A3. The overall purpose of this paper is to bring these techniques to the attention of statisticians working in the design of chemical experiments.

In Section 2 we introduce the model of the general consecutive reaction which is used to exemplify the design procedure. A brief summary of optimum design theory is in §3, with the D- and D_s -optimum designs in §4. We then extend the theory and designs to improve their applicability. In §5 we use compound designs to enable the experimenter to make an informed choice between designs for estimating order and those for rate. In §6 we find a Bayesian design, appropriate when there is appreciable prior uncertainty about the orders of the two reactions. In §7 we explain the numerical procedures used for our calculations. The paper closes in §8 with a discussion of related work and some suggestions for further developments. The direct method for finding derivatives is described in the Appendix.

2 Rates and Orders of Reaction

2.1 The General Consecutive Reaction

The example given by Valko and Vajda (1984) is of a complicated series of reactions from combustion chemistry. Our description of the method in the Appendix covers the calculation of sensitivities for general sets of equations. In the body of the paper we consider relatively simple examples which however illustrate the need for the numerical calculation of sensitivities in experimental design.

The model for the two consecutive first-order reactions considered by Box and Lucas (1959) is



The concentrations of any, or all, of A , B and C can be measured. Since both reactions are first order, the kinetic differential equations for $[A]$, $[B]$ and $[C]$, the concentrations of the chemical compounds A , B and C as functions of

time t are

$$\begin{aligned}\frac{d[A]}{dt} &= -\theta_1[A] \\ \frac{d[B]}{dt} &= \theta_1[A] - \theta_2[B] \\ \frac{d[C]}{dt} &= \theta_2[B].\end{aligned}\tag{2}$$

Given the initial concentrations of A , B and C , an explicit algebraic solution can be found for the concentrations as a function of time. The analytical derivatives of this function with respect to the parameters are then used in the design of experiments.

The equations (2) are a straightforward example of the sets of linear differential equations that arise from kinetic models. One instance (Atkins 1998, p.794) is the consecutive radioactive decay of Uranium 239 through Neptunium to Plutonium. Reactions are rarely so simple. More complicated sets of first-order equations cannot always be solved analytically - examples are given by Jennrich and Bright (1976), in Chapter 5 of Bates and Watts (1988), in Chapter 8 of Seber and Wild (1989) and by Nathanson and Saidel (1985).

When the chemical reactions are not first-order we typically obtain non-linear equations. In the general consecutive reaction the equations (2) are replaced by

$$\begin{aligned}\frac{d[A]}{dt} &= -\theta_1[A]^{\lambda_1} \\ \frac{d[B]}{dt} &= \theta_1[A]^{\lambda_1} - \theta_2[B]^{\lambda_2} \\ \frac{d[C]}{dt} &= \theta_2[B]^{\lambda_2}.\end{aligned}\tag{3}$$

Interest is in estimation of the orders $\lambda = (\lambda_1, \lambda_2)$ as well, perhaps, as of the rates $\theta = (\theta_1, \theta_2)$. If both λ_1 and λ_2 are equal to 1, the reactions are first-order and (3) reduces to (2). Chemical theory shows that the orders will be small integers, or ratios of small integers. Experimental results yielding other orders are often an indication that some parts of the reaction scheme have not been correctly modelled. For example, the first reaction in (1) might be reversible, the model being



with θ_3 the rate of the reverse reaction. If the model for consecutive reactions (1) is fitted when (4) holds, the value of λ_1 may be far from a ratio of small integers. Such elaborations of reaction schemes are frequent in combustion and environmental studies. (Atkins 1998, p.778) describes a series of increasingly complicated reaction schemes for the decomposition of the atmospheric pollutant ozone. Experiments to estimate the order of reactions are therefore important in such model building.

The main example in our paper is the general consecutive reaction with kinetic equations (3). The first of the three equations can be solved analytically to give the concentration of chemical A at time t as

$$[A] = \{1 - (1 - \lambda_1)\theta_1 t\}^{1/(1-\lambda_1)} \quad (\lambda_1, \theta_1, t \geq 0; \lambda_1 \neq 1), \quad (5)$$

if it is assumed that the initial concentration of A is 1. If λ_1 is less than one, the reaction is complete at $t = 1/\theta_1$, after which $[A] = 0$. If (5) is of main interest and the concentration $[A]$ is measured, the reaction is that of general decay, reducing to exponential decay when $\lambda_1 = 1$. However, when both reactions are of interest, substitution of $[A]$ from (5) into the expression for $[B]$ (3) yields a differential equation from which analytical expressions for $[B]$ can only be found for a few values of λ_1 and λ_2 . Even if an explicit expression is available for the concentration for these particular values, the expression will not provide the derivatives of the concentration with respect to the orders λ_1 and λ_2 which we require in order to design the experiment.

In the chemical literature the derivatives of the response with respect to the parameters are called “first-order sensitivities”. Their importance in model building is stressed, for example, by Pilling and Seakins (1995, p.203). We find numerical values of the sensitivities by solving a further set of differential equations. Solution of the equations provides a grid of numerical values of the sensitivities, which have to be interpolated in the construction of the design. These methods are described in the Appendix and §7.

3 Experimental Design Theory

3.1 Notation and Methods

In this section we establish notation, recall the standard results of the theory of optimum experimental design and introduce the direct method of finding parameter sensitivities which are necessary for calculating optimum designs. The next subsections describe D-optimum designs, appropriate if all the parameters, both rates and orders of reaction, are of equal interest and

D_s -optimum designs for only a subset of the parameters. In our example we shall be interested in orders of reaction, with the rates perhaps treated as nuisance parameters.

The experiment consists of measuring the concentration of B after the reaction has been running for a time t . One experimental run yields one observation y_i and the experimental design is a list of the n times, $t_i, i = 1, \dots, n$, not necessarily distinct, at which measurements are to be made. Comments on experimental design when observations can be made at several time points during one run are given in §8.

For mathematical convenience we consider the more general definition of an experimental design ξ which is a continuous measure specifying a set of p distinct points in a design region \mathcal{T} and the proportions, w_i , of observations taken at these points

$$\xi = \left\{ \begin{array}{l} t_1, \dots, t_p \\ w_1, \dots, w_p \end{array} \right\}, \quad 0 < w_i \leq 1, \quad \sum_{i=1}^p w_i = 1.$$

The design region is determined by the technical consideration of the times at which it is possible to take measurements. The times t_i are the points of support of the design ξ and w_i the design weights. In practice, when only n observations can be taken, an exact design will be required. Often the optimum exact design is approximated by a design with the number of trials at t_i the integer closest to nw_i .

We assume that the concentration of only one chemical is measured. The extension to multiple responses is presented in §3.4. The response, y , for the nonlinear regression model has mean

$$E(y) = \eta(t, \psi),$$

and dispersion matrix $D(y) = \sigma^2 I, \sigma^2 > 0$ and I the $n \times n$ identity matrix. Here η is a univariate function nonlinear in at least one of the k parameters $\psi = (\psi_1, \dots, \psi_k)$ and $t \in \mathcal{T}$ denotes time. We assume that random errors are additive and that they are independently, identically, normally distributed with zero mean and constant variance σ^2 . The appropriate method of analysis of the results is nonlinear least squares.

Most of the criteria in optimal design are functions of the information matrix for the set of parameters of interest. For linear models the inverse of this matrix is the variance-covariance matrix of the parameter estimator. In a non-linear model this property holds asymptotically. The information matrix of a design ξ for the k parameters ψ is given by

$$M(\xi, \psi) = F^T W F,$$

where

$$F = \begin{pmatrix} f^T(t_1, \psi) \\ \vdots \\ f^T(t_p, \psi) \end{pmatrix}$$

is a $p \times k$ matrix and the i -th row vector $f^T(t_i, \psi)$ has j -th element

$$f_j(t_i, \psi) = \frac{\partial \eta(t_i, \psi)}{\partial \psi_j}, \quad \text{for } j = 1, \dots, k,$$

called the parameter sensitivity, and

$$W = \text{diag}\{w_1, \dots, w_p\}.$$

The information matrix thus depends on the unknown parameters, ψ . In §4 and §5 we adopt a best guess ψ^o for the parameters and consider designs which maximize an appropriate function of $M(\xi, \psi)$ evaluated at $\psi = \psi^o$ (Chernoff 1953). Such designs are termed locally optimum. The Bayesian designs of §6 require a prior distribution of the parameters.

Expressions for the required derivatives for our example are given in Section A.2 of the Appendix. They come out as a special case of the direct method for the univariate expected response. The method for a multivariate response is presented in section A.1 of the Appendix. Another special case for the more complicated reaction (4) is given in A.3.

For the general consecutive reaction the vector of parameters is

$$\psi = (\theta_1, \theta_2, \lambda_1, \lambda_2)$$

and so the sensitivities f_1 and f_2 are for θ_1 and θ_2 , with f_3 and f_4 for λ_1 and λ_2 . The sets of differential equations formed by (A6) and each of (A7) are nonlinear. They will only be linear if the orders of reaction are known and equal to one, when only f_1 and f_2 need to be found. With such a set of linear equations the methods described by Bates and Watts (1988) and Seber and Wild (1989) for the solution of sets of linear equations can be used. We used Maple V, Release 4, to solve four sets of two nonlinear equations.

In order to design experiments we need either prior values for the parameters or a prior distribution. In most of the paper we assume point prior orders of reaction $\lambda^o = (\lambda_1^o, \lambda_2^o)$ and prior rates of reaction $\theta^o = (\theta_1^o, \theta_2^o)$. We begin with the solution of the differential equations for the four combinations of prior orders of reaction in which both λ_1^o and λ_2^o took values 1 and 2. In all cases the rates of reaction were taken as $\theta_1^o = 0.7$ and $\theta_2^o = 0.2$. The initial concentration of A was one, the other two concentrations initially being zero. Fig.1 shows the response $[B]$ as a function of time for the four sets of λ values.

These curves fall into two groups - those with $\lambda_2 = 2$ decrease very slowly from the maximum yield compared to those with $\lambda_2 = 1$. So experiments at high values of t may be expected to be required when $\lambda_2 = 2$. Fig.2 shows the four derivatives $f_j(t, \psi)$ when $\lambda^o = (1, 1)$ and $\theta^o = (0.7, 0.2)$, which can be compared with Fig. 1 of Box and Lucas (1959). However, because Box and Lucas were assuming that the reactions were known to be first order, they only show the two derivatives with respect to the rates θ .

3.2 D-optimum Designs

Since our interest is in precise estimation of the parameters in our model, it is appropriate to find D -optimum designs. These maximize the logarithm of the determinant of the information matrix, $\log |M(\xi, \psi)|$ or, equivalently, minimize the asymptotic generalized variance of the parameter estimator, that is the determinant of its dispersion matrix. However, in a non-linear model this very useful property requires large numbers of observations which makes the criterion less attractive. Alternative criteria which take into account the parameter and intrinsic curvatures of the model, for example the criterion introduced by Hamilton and Watts (1985), involve complicated three dimensional matrices of second-order sensitivities and so are an order of magnitude more difficult to calculate than the designs found here. Furthermore, there is no theoretically based way of checking that the calculated design is indeed optimum, a serious drawback where the results are obtained by numerical rather than analytical calculations. On the other hand, if the linear approximation of the model is accurate enough, so that the curvatures can be ignored, the D -optimality criterion is not only much easier to apply but also, due to the Equivalence Theorem of Kiefer and Wolfowitz (1960) it is possible to check the optimality of a candidate continuous D -optimum design.

This equivalence theorem relates maximization of the determinant of the information matrix to minimization of the maximum variance of the predicted response over \mathcal{T} . With the standardized variance of the prediction at t defined by

$$d(t, \xi, \psi) = f^T(t, \psi)M^{-1}(\xi, \psi)f(t, \psi), \quad (6)$$

the relevant part of the Equivalence Theorem states that, for the optimum design, ξ^* , the maximum value of $d(t, \xi^*, \psi)$ over the design region, \mathcal{T} , is k , the number of parameters in the model, and further that this maximum value is attained at the support points t_i^* of ξ^* . The theorem also provides a basis for algorithmic construction of locally D -optimum designs (Wynn 1970).

It follows from Carathéodory's Theorem (Silvey 1980, p.72) that a continuous D -optimum design is based on at most $k(k + 1)/2$ design points.

Often locally D -optimum designs for nonlinear models are based on exactly k points of support, when the weights associated with the support points are equal to $1/k$ (Silvey 1980, p.42).

3.3 D_s -optimum Designs

If only a subset of s of the parameters, $\psi_{(2)}$, is of interest, let the parameters be partitioned as $\psi = (\psi_{(1)}, \psi_{(2)})$ with the information matrix $M(\xi, \psi)$ partitioned so that the information for $\psi_{(1)}$ is $M_{11}(\xi, \psi)$. Then the D_s -optimum design for $\psi_{(2)}$ maximizes $\log\{|M(\xi, \psi)|/|M_{11}(\xi, \psi)|\}$. The Equivalence Theorem for D_s -optimum designs states that, for the optimum measure ξ^* , the analogue of the standardised variance of prediction is

$$d(t, \xi^*, \psi) = f^T(t, \psi)M^{-1}(\xi^*, \psi)f(t, \psi) - f_{(1)}^T(t, \psi)M_{11}^{-1}(\xi^*, \psi)f_{(1)}(t, \psi) \leq s, \quad (7)$$

where $f_{(1)}^T(t, \psi)$ is a vector of sensitivities for $k - s$ parameters. An early example of the use of D_s -optimum designs in the study of kinetic models is Hunter, Hill, and Henson (1969).

3.4 Multivariate Response

Suppose that the concentration of more than one chemical is measured. There will then be a model for each expected response giving a matrix F_i of parameter sensitivities for the i th response, $i = 1, \dots, m$,

$$F_i = \{f_{ij}(t_u, \psi)\} = \left\{ \frac{\partial \eta_i(t_u, \psi)}{\partial \psi_j} \right\},$$

where $u = 1, \dots, p$ denotes the design points and $j = 1, \dots, k$ denotes the parameters. The generalization of the single-response case is that now the observations follow the model

$$y_{iu} = \eta_i(t_u, \psi) + \varepsilon_{iu},$$

with

$$E(\varepsilon_{iu}) = 0, \quad E(\varepsilon_{iu}\varepsilon_{lv}) = \begin{cases} 0 & \text{if } u \neq v \\ \sigma_{il} & \text{if } u = v \end{cases},$$

when the variance-covariance matrix of the responses is

$$\Sigma = \{\sigma_{il}\}_{i,l=1,\dots,m}.$$

Draper and Hunter (1966), following arguments similar to those of Box and Lucas (1959) for the single-response case, show that for normally distributed errors the information matrix is given by

$$M(\xi, \psi) = \sum_{i=1}^m \sum_{l=1}^m \sigma^{il} F_i^T W F_l, \quad (8)$$

where $\Sigma^{-1} = \{\sigma^{il}\}_{i,l=1,\dots,m}$. We can calculate all necessary sensitivities using the direct method presented in the Appendix.

As an example we give in Appendix A3 the calculation of the sensitivities for model (4) when all three concentrations are measured. With six parameters and three responses there are in all 18 sensitivities. Although we do not calculate designs, the locally D-optimum design would, as before, maximize $|M(\xi, \psi^0)|$.

4 Optimum Designs for the Generalised Consecutive Reaction

We now use the preceding theory to calculate optimum designs for a variety of relative interests in the four parameters of the model. We start with *D*-optimum designs which are appropriate if all parameters are of equal interest. We then consider designs for the rates of reaction under the assumption that the orders of reaction are known. These designs generalize those of Box and Lucas (1959) to values of λ_1 and λ_2 which are not both one. The section concludes with designs for precise estimation of the orders, when the rates are unknown. The next section combines these criteria in a parameterised compound criterion. Varying the parameter can lead to a design which provides a balance between the three criteria of this section.

4.1 D-optimum Designs for Rate and Order of Reaction

If both rate of reaction and the order are of interest, the relevant design criterion is that of D-optimality, §3.2, maximizing $\log |M(\xi, \psi^o)|$.

For our four-parameter model the D-optimum design has four points of support, with weight 0.25 at each. However, the values of the design points depend on the prior values of the parameters. Table 1 gives the optimum design points for the four values of λ^o : (1,1), (1,2), (2,1) and (2,2) and a constant prior for θ , $\theta^o = (0.7, 0.2)$. The design region is $\mathcal{T} = [0,50]$. These

Prior orders of reaction (λ_1^o, λ_2^o)	time			
	t_1	t_2	t_3	t_4
(1,1)	0.80	2.85	7.05	15.90
(2,1)	0.51	2.36	7.30	18.26
(1,2)	0.83	2.91	8.05	40.39
(2,2)	0.57	2.65	9.68	50.00

Table 1: D-optimum designs for both rate and order. The weights are 0.25 at each design point.

designs were found by searching over the four continuous values of time, but with the weights held known at 0.25.

This table shows that the two lower design points depend principally on the prior value of λ_1 . However, the most spectacular effect is the dependence of t_4 on the prior value of λ_2 . When $\lambda_2^o = 2$ both experimental designs have values of t_4 above 40, whereas for $\lambda_2^o = 1$ the values are below 20. Fig.1 shows the relationship between these designs and the values of $[B]$. For $\lambda_2 = 1$, the concentration is negligible above $t = 40$, but is still appreciable in this region if $\lambda_2 = 2$.

The plots of Figures 3 and 4 show the usefulness of the variance $d(t, \xi^*, \psi)$ in checking optimum designs and exploring their structure. Fig.3 shows the variance function for $\lambda^o = (1, 1)$. There are four maxima, one at each optimum design point, at each of which $d(t, \xi^*, \psi) = 4$. The same is true for the D-optimum design for $\lambda^o = (2, 2)$ in Fig.4 except that now the fourth design point is on the boundary of the design region and the derivative of the variance is not zero. If the design region were increased in size a design with a larger value of t_4 would be obtained, which would have an increased value of the design criterion. The designs were constructed with weights w_i taken as 0.25. Since the designs satisfy the conditions for an optimum design, the figures also confirm that the weights for the optimum designs are indeed 0.25.

4.2 D-optimum Designs for the Rate of Reaction

If the orders of reaction λ_1 and λ_2 are known, it makes sense to find the D-optimum designs for the rates θ_1 and θ_2 . These designs, maximising $\log |M_{11}(\xi, \theta^o)|$, have two design points, with weight 0.5 at each. They are listed in Table 2, the optimum points when $\lambda = (1, 1)$ being originally calculated by Box and Lucas (1959).

It is surprising that the designs depend so little on the assumed values

Orders of reaction $\lambda = (\lambda_1, \lambda_2)$	Times	
	t_1	t_2
(1,1)	1.23	6.85
(2,1)	1.01	7.70
(1,2)	1.19	7.52
(2,2)	1.06	10.09

Table 2: D-optimum designs for rate, taking prior $\theta^o = (0.7, 0.2)$ when orders are assumed known. We take four values of λ_1 and λ_2 . The weights now are 0.5 at each design point.

of λ_1 and λ_2 . Comparison of the values of t_2 in Table 2 with those of t_3 in Table 1 show surprising agreement. The large values of time in Table 1 are not present in the optimum designs for the rate only.

4.3 D_s -optimum Designs for the Orders of Reaction

In model building determination of the orders of the reactions, that is λ_1 and λ_2 , will often be of primary importance, with the values of the rate constants θ_1 and θ_2 being of secondary importance. In such situations the D_s -optimum design for the λ 's is appropriate. To find this D_s -optimum design we partition the information matrix $M(\xi, \psi)$ so that the information matrix for θ_1 and θ_2 is $M_{11}(\xi, \psi)$. The required design then maximizes $\log\{|M(\xi, \psi)|/|M_{11}(\xi, \psi)|\}$.

The D_s -optimum designs are given in Table 3, which is to be compared with Table 1. As before, the designs fall into two broad classes: those with $\lambda_2^o = 2$ both have part of the design effort at $t = 50$, the maximum value of the experimental region \mathcal{T} , whereas, for $\lambda_2^o = 1$, the maximum time is less than half this value. Although the times at which measurements are taken are broadly similar in Tables 1 and 3 for each pair of values of λ^o , the times in the D_s -optimum designs have a larger range than those for the D-optimum criterion. In all four cases the design weights are unequal, with about 0.4 on the maximum time point. More experimental effort is focused on the large value of time, which leads to more precise estimation of λ rather than of θ . All designs, of course, satisfy the equivalence theorem (7) with $s = 2$.

5 Compound Optimum Designs

Each of the three designs of the previous section is tailor-made for solving one aspect of the design problem. We now consider the use of compound

Prior orders of reaction (λ_1^o, λ_2^o)	time and weights			
	t_1	t_2	t_3	t_4
	w_1	w_2	w_3	w_4
(1,1)	0.54	3.13	7.48	17.61
	0.16	0.25	0.18	0.41
(2,1)	0.36	2.57	7.49	20.91
	0.22	0.22	0.17	0.39
(1,2)	0.55	3.15	8.57	50.00
	0.14	0.26	0.18	0.42
(2,2)	0.40	2.93	9.49	50.00
	0.21	0.24	0.18	0.37

Table 3: D_s -optimum designs for order for four prior values of λ_1 and λ_2 , with $\theta^o = (0.7, 0.2)$. Both weights and design points have to be found numerically.

optimum designs by which the experimenter can find a single design which strikes a balance between the three objectives.

Each of the three design criteria we have considered is an example of a convex design criterion. Linear combinations of such criteria with non-negative weights will also be convex design criteria. The compound design criterion used here is an example of such a linear combination, for which designs are found to maximize

$$\begin{aligned} \Phi(\xi, \psi) &= (1 - \alpha) \log |M_{11}(\xi, \psi)| + \alpha \log \{|M(\xi, \psi)| / |M_{11}(\xi, \psi)|\} \\ &= (1 - 2\alpha) \log |M_{11}(\xi, \psi)| + \alpha \log |M(\xi, \psi)|. \end{aligned} \quad (9)$$

In (9) α ($0 \leq \alpha \leq 1$) expresses the experimenter's relative interest in determination of the parameters of the reaction, with $\alpha = 1$ corresponding to interest solely in order determination. When $\alpha = 0.5$ the criterion becomes a multiple of that for D-optimality for both orders and rates. Such multiplication of the design criterion does not affect the optimum design. Finally, when $\alpha = 0$, the criterion becomes that of D-optimality when it is assumed that all orders of reaction are known. Cook and Wong (1994) give an interpretation of compound designs as satisfying a constrained optimisation. An example of compound designs and of the use of the equivalence theorem is given by Atkinson and Bogacka (1997) for their two-parameter model. Here

we extend the equivalence theorem to models with an arbitrary number of parameters.

Let the k parameters be partitioned as in §3.3, that is $\psi = (\psi_{(1)}, \psi_{(2)})$, where $\psi_{(2)}$ is the s -dimensional vector of the parameters of primary interest. Then $\psi_{(1)}$ is of dimension $k - s$. It is convenient to write this dimension as $k - s = r$. We again used the standardised variance of prediction at time t both in the construction of designs and to check the optimality. The compound criterion (9) is a weighted linear combination of D- and D_s -criteria with weights $1 - \alpha$ and α . The variance function is then the weighted linear combination of the variances for the individual criteria with the same weights. Therefore the optimum design ξ_c^* is such that

$$\begin{aligned} d_c(t, \xi_c^*, \psi) &= (1 - 2\alpha)\{f_{(1)}^T(t, \psi)M_{11}^{-1}(\xi_c^*, \psi)f_{(1)}(t, \psi)\} \\ &+ \alpha\{f^T(t, \psi)M^{-1}(\xi_c^*, \psi)f(t, \psi)\} \\ &\leq (1 - 2\alpha)r + \alpha(r + s) \\ &= r + \alpha(s - r), \end{aligned} \tag{10}$$

where $r = k - s$. The bound on the variance then depends upon α unless $s = r = k/2$.

In many kinetic models there are fewer rate constants than orders of reaction, so we may have $r < s$. Simple examples of such models are in Pilling and Seakins (1995, Chapter 1). But in our example of the general consecutive reaction, $r = s = 2$, so that the variance in (10) does not depend on α .

It is unlikely that an experimenter will be able to specify a value of α for use in the compound criterion. Instead we find designs maximizing the compound criterion for a series of values of α . For each we calculate the efficiencies of the design relative to the D-optimum design for the estimation of rate, the D_s -optimum design for estimation of order and relative to the D-optimum design for both. A plot of these efficiencies against α makes it possible to choose a design with a balance of efficiencies for all aspects of the problem.

The number of designs to be considered has increased, as there will be a set of compound designs, depending on α , for each pair of values $(\lambda_1^o, \lambda_2^o)$. We only illustrate the designs for $\lambda_1^o = \lambda_2^o = 1$. The general structure of designs for other prior orders of reaction is similar.

Fig.5 gives a plot of the four time points of the compound designs as a function of α . For $\alpha = 1$ these times are those of the D_s -optimum design of Table 3. For $\alpha = 0.5$ we have the D-optimum design of Table 1 and, as $\alpha \rightarrow 0$, the design approaches that of the D-optimum design of Table 2, which assumes that the orders of reaction are known. The times of experimentation

do not change appreciably with α . But the changes in weights, shown in Fig.6, are both more dramatic and more informative about the changing structures of the designs. When $\alpha = 1$ we have the design of Table 3 in which the maximum design weight is 0.41. When $\alpha = 0.5$, the D-optimum design for all four parameters is obtained in which all weights are one quarter. As α approaches zero two of the weights tend to zero and two, w_1 and w_3 , tend to one half, the values obtained for $\alpha = 0$.

These figures show the behaviour of the compound designs as α changes. To choose a value of α which yields a design reflecting the experimenter's interests, requires calculation of the efficiency of a proposed design for the three specific aspects of interest. The efficiencies are calculated using the optimum designs for these aspects. Let the D-optimum design for estimating θ be ξ_θ^* . Then the efficiency of the compound design if only θ is of interest is

$$E_\theta = 100\{|M_{11}(\xi_c^*, \psi^o)|/|M_{11}(\xi_\theta^*, \psi^o)|\}^{1/r}, \quad (11)$$

the $(1/r)$ th power of the determinantal ratio being required to provide an efficiency in terms of the variance of individual parameters and so of numbers of trials. Likewise, if the D_s -optimum design for estimating λ_1 and λ_2 is ξ_λ^* , the relevant efficiency is

$$E_\lambda = 100 \left\{ \frac{|M(\xi_c^*, \psi^o)|/|M_{11}(\xi_c^*, \psi^o)|}{|M(\xi_\lambda^*, \psi^o)|/|M_{11}(\xi_\lambda^*, \psi^o)|} \right\}^{1/s}. \quad (12)$$

Finally, if the D-optimum design for θ and λ is ξ_D^* , the efficiency is

$$E_D = 100\{|M(\xi_c^*, \psi^o)|/|M(\xi_D^*, \psi^o)|\}^{1/k}. \quad (13)$$

A plot of these three efficiencies against α is given in Fig.7. At the borders of the range of α , the compound design is good for only one of the aspects of the problem: either estimation of the rates of reaction, when α is close to zero, or, when α is close to one, the estimation of orders with rates as nuisance parameters. When $\alpha = 0.5$ the compound design is 100% efficient for estimation of both sets of parameters: it is D-optimum for θ and λ . An interesting choice of α is 0.73 where the curves for E_D and E_λ intersect and the efficiencies are approximately 96%. Such a design, along with any in the range $0.5 \leq \alpha \leq 1$, has good efficiency for both determination of the orders of reaction and estimation of the rate parameters.

6 Bayesian Designs

6.1 Theory

The compound designs of the previous section were illustrated for both unknown orders of reaction taken to have prior values one. However if the prior information about possible orders of reaction is much less precise, it may be important to find designs which take into account several values of λ_1 and λ_2 in one experiment. Accordingly, we now consider Bayesian designs which incorporate prior information about parameters in the form of distributions, rather than as point estimates.

In the customary form of Bayesian optimum design a prior distribution is assumed for the unknown parameters and design criteria like those of §3 are replaced by their expectations over the prior. This form of Bayesian experimental design is described by Chaloner and Larntz (1989) for logistic models for binary data. Chaloner and Verdinelli (1995) give a full review. Several examples for nonlinear models are given by Atkinson and Donev (1992), Chapter 19. Atkinson and Bogacka (1997) give an example for the order of the general decay model. We extend this to our multiparameter model and demonstrate the importance of the related equivalence theorem in checking optimum designs.

6.2 Bayesian D_s Optimum Designs for λ

The effect of prior uncertainty about λ is illustrated by extending the D_s -optimum design criterion for the order of reaction of §§3.3 and 4.3 to a discrete prior distribution for λ with equal weight put on each pair $(\lambda_1^j, \lambda_2^j)$. We consider the four pairs: (1, 1), (1, 2), (2, 1) and (2, 2). Then the criterion to be maximized is

$$\Phi(\xi) = \sum_{j=1}^4 0.25 \log\{|M(\xi, \theta^o, \lambda^j)|/M_{11}(\xi, \theta^o, \lambda^j)\}, \quad (14)$$

where $\lambda^j = (\lambda_1^j, \lambda_2^j)$ and $\theta^o = (\theta_1^o, \theta_2^o)$. The optimum Bayesian design ξ_B^* will be such that

$$\begin{aligned} d(t, \xi_B^*, \theta^o, \lambda) &= 0.25 \sum_{j=1}^4 \{f^T(t, \theta^o, \lambda^j) M^{-1}(\xi_B^*, \theta^o, \lambda^j) f(t, \theta^o, \lambda^j) \\ &\quad - f_{(1)}^T(t, \theta^o, \lambda^j) M_{11}^{-1}(\xi_B^*, \theta^o, \lambda^j) f_{(1)}(t, \theta^o, \lambda^j)\} \leq s. \end{aligned} \quad (15)$$

The locally D_s -optimum designs of Table 3 put trials at k design points, in our case four. It is a feature of Bayesian designs that, as prior uncertainty

about the parameters increases, so does the number of design points. As Table 4 shows, the optimum design maximizing (14) has five, rather than four, points. That the design is indeed optimum is shown (in Fig. 8) by the plot of the variance (15) which attains its maximum value of two at these five design points, one of which is the end of the experimental region. The general structure of this design is arguably not surprising when the design is considered as an averaging of those of Table 3. All designs have trials around 0.5, 3 and 8. The designs for $\lambda_2^o = 1$ have the fourth point around $t = 20$ whereas, for $\lambda_2^o = 2$ the fourth time is at the maximum of the experimental region. Of course, the five-point design is not an exact average of the other designs, these comments serving only to clarify the structure of the design.

Finally, to stress the importance of the equivalence theorem in checking proposed designs, Table 4 also includes a four-point Bayesian design. This is the best design we were able to find with support on four points. Relative to the five-point design it has an efficiency of $100(4.753/3.548)^{1/2} = 86.4\%$. As the plot of Fig.9 shows, the design can be improved by including experiments at times higher than 22.2 and that 50 is indicated. However, including 50 in a four-point design leads to designs with a variance function which is above two at other points of the design region. These designs have a smaller determinant than the design of Table 4 and so are not as good. The design in the table is the best we could find by numerical search from many starting points, guided by plots like Fig.9. However, since the equivalence theorem is not satisfied, we cannot be sure that a better four-point design does not exist. We can however be sure that the five-point design is the optimum Bayesian design for the four-valued prior we have used. In addition, the five point design allows for checking for lack of fit of the four parameter model. For example, if in fact the reaction were



experiments at high values of t from a five-point design would be informative about the failure of the irreversible scheme (1) which is the subject of this paper.

7 Numerical Procedures

7.1 Interpolation

The direct method described in the Appendix is a helpful and elegant tool for finding the parameter sensitivities in all cases of nonlinear models where the expected response (uni- or multivariate) is given by the solution of a

	Five-point Design					
Time	0.45	2.96	7.62	19.55	50.0	0.4753E-22
Weight	0.18	0.24	0.17	0.27	0.14	
Determinant						
	Four-point Design					
Time	0.45	2.96	7.78	22.20		0.3548E-22
Weight	0.19	0.25	0.20	0.36		
Determinant						

Table 4: Bayesian D_s -optimum design for λ_1 and λ_2 with prior putting weights 0.25 on the orders of reaction (1,1),(1,2),(2,1) and (2,2); $\theta^o = (0.7, 0.2)$.

set of differential equations. In most situations the sets of equations have to be solved numerically and we found the Maple procedure “`dsolve`” very fast. The numerical procedure for the solution of sets of differential equations produces answers at a grid of time points. We have used a grid with intervals of 0.1. As Figure 2 shows, the plots of sensitivities against time do not vary rapidly, so that the choice of this interval is not critical. In order to design experiments we need the values of the derivatives at any real t within the experimental region and so have to interpolate the numerical solutions from Maple. We used Fortran interpolation routines from Press et al. (1992, Chapter 3). Two procedures are required. The subroutine `locate` (Press et al., §3.4) finds, by bisection, the nearest entry in the table of derivatives to the value of t for which the derivatives are required. Subroutine `polint` (Press et al., §3.1) uses differences to fit a cubic polynomial interpolation formula.

The design of experiments requires much less computation than fitting the models to kinetic data. In experimental design we have a fixed set of parameter values for which the numerical solution of the differential equations is performed once. These tabulated values are then interpolated for each t required in the optimization of the design criterion. In parameter estimation a search is also performed, now to minimise the residual sum of squares, each time for a different set of parameter values, so that the numerical solution of the equations has to be repeated.

In the example in this paper we did not find computing time to be a constraint. If it were, some time might be saved by replacing the subroutine `locate` by direct calculation of the required index in the table of sensitivities, using the information that the table is based on equal increments of time.

However the method we used has the advantage of flexibility for use in tables of sensitivities not necessarily calculated at equal time intervals.

7.2 Design

Since we are finding designs defined by a measure we only have to find the optimum of a convex function. If the design weights are not known this is a function of seven variables - four design points and four non-negative weights which sum to one. We used a standard conjugate gradient algorithm with the transformations in Atkinson and Donev (1992, p.104) to ensure that the weights were non-negative and summed to one. In all cases we used the equivalence theorem to check that we had indeed found the optimum design and that it contained the correct number of points.

8 Discussion

In this paper we have used the “direct” method to find the parameter sensitivities needed to design optimum experiments for the orders of chemical reactions described by systems of differential equations. We have assumed that the experimental design problem is to choose times at which measurements are to be taken. A similar strategy is followed by Nathanson and Saidel (1985) who use the direct method in evaluating two design criteria for a ten-parameter compartmental model for human iron exchange which yields a set of linear differential equations. This work appears unknown to most statisticians. Our extensions here are to nonlinear equations and to the use of the equivalence theorem and compound designs in investigating numerous properties of the designs and their dependence on the precise questions being asked. A question we have not investigated is the dependence of the designs on the prior values of θ . Some indication of the dependence of kinetic models on changes in rate constants can be gleaned from the examples in Chapter 18 of Atkinson and Donev (1992): in general, designs for kinetic models are more sensitive to changes in reaction orders than they are to changes in rate.

We have assumed throughout that we have additive errors of constant variance. As a referee has commented, chemical measurements are often heteroscedastic. In analytical chemistry, Horwitz’ rule (Horwitz 1982) suggests a linear relationship between log standard deviation and log expected value of the response. The variance can therefore be stabilised by a power transformation. Since the response and time are related by a kinetic model, both sides of models such as (5) have to be transformed before they become statistical models by the addition of additive homoscedastic error. Horwitz’

rule indicates a transformation within the parametric family of power transformations (Box and Cox 1964) with $\lambda = 0.14$. Since $\lambda = 1$ corresponds to no transformation and $\lambda = 0$ to the log transformation, this law makes very strong assumptions about the accuracy of measurements at low concentrations. The effect on designs for kinetic models as λ varies in this family is investigated by Atkinson (2002), who shows that choice of λ can have an appreciable effect on the optimum design. As λ approaches zero, measurements are increasingly made at lower concentrations. However, in a paper on the analysis of data from analytical chemistry, Rocke and Lorenzato (1995), particularly Appendix A, argue that Horwitz' rule overstates the accuracy of measurements at low concentrations when a single measuring instrument is used, and that a two-component error model is more appropriate. Preliminary investigations suggest that this error model yields designs which allow for heteroscedasticity without overemphasising measurements at low concentrations.

If several measurements are to be taken from a single experimental run, allowance should be made for the correlation structure of the observations over time. However, for many chemical experiments the choice of sampling time is not an issue. Lutchen and Saidel (1982) compare ten designs for a dynamic model of lung function when a fixed number of equally spaced observations are to be taken. Under such a measurement scheme, the experimental design specifies the initial conditions and the variation of factors during the experiment.

The design of experiments for chemical plants is more complicated than that described here, although similar principles apply. The complications are that energy and mass balances have to be included in the differential equations forming the model. Reactants can be introduced and product removed during the course of the reaction. During this process the temperature may change, with consequent changes in reaction rates. A full description of these problems in the context of chemical engineering and of the numerical techniques needed for their solution is given by Bauer, Bock, Körkel, and Schlöder (2000). They do however assume heteroscedastic errors. For the simpler problem treated here, standard packages for the solution of differential equations and optimization suffice.

Acknowledgements

We are grateful to Luc Pronzato and Weng Kee Wong for important references.

A Appendix

A.1 The Direct Method

This appendix is intended to provide adequate guidance in the use of the direct method in a variety of statistical problems. We first describe the method in general and then give examples of its application to the kinetic models (1) and (4) of Section 2.

Let $\eta(t, \psi) = (\eta_1(t, \psi), \dots, \eta_m(t, \psi))^T$ be a set of m functions representing the expected multivariate response at time $t \in T$ with common parameters $\psi = (\psi_1, \dots, \psi_k)^T$. The functions η_i , $i = 1, \dots, m$, are solutions of a set of differential equations

$$\begin{aligned} \frac{d\eta_1(t, \psi)}{dt} &= g_1(\eta, t, \psi) \\ &\vdots \\ \frac{d\eta_m(t, \psi)}{dt} &= g_m(\eta, t, \psi) \end{aligned} \quad (\text{A1})$$

with initial conditions given by

$$\eta(t_0, \psi) = \eta_0(\psi), \quad (\text{A2})$$

where $g_i(\cdot)$, $i = 1, \dots, m$, are known differentiable functions of η , t and ψ . Furthermore, let f_j be the vector of sensitivities for the j th parameter, that is the vector of derivatives of the response functions with respect to ψ_j :

$$f_j(t, \psi) = (f_{1j}(t, \psi), \dots, f_{mj}(t, \psi))^T = \left(\frac{\partial \eta_1(t, \psi)}{\partial \psi_j}, \dots, \frac{\partial \eta_m(t, \psi)}{\partial \psi_j} \right)^T, \quad (\text{A3})$$

$$j = 1, \dots, k.$$

In matrix notation the sensitivity equations may be written as

$$\frac{df_j(t, \psi)}{dt} = \frac{\partial g(\eta, t, \psi)}{\partial \eta} f_j(t, \psi) + \frac{\partial g(\eta, t, \psi)}{\partial \psi_j}, \quad j = 1, \dots, k, \quad (\text{A4})$$

or, in full notation, we get the sets of equations

$$\begin{bmatrix} \frac{df_{1j}(t, \psi)}{dt} \\ \vdots \\ \frac{df_{mj}(t, \psi)}{dt} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1(\eta, t, \psi)}{\partial \eta_1} & \cdots & \frac{\partial g_1(\eta, t, \psi)}{\partial \eta_m} \\ \vdots & & \vdots \\ \frac{\partial g_m(\eta, t, \psi)}{\partial \eta_1} & \cdots & \frac{\partial g_m(\eta, t, \psi)}{\partial \eta_m} \end{bmatrix} \begin{bmatrix} f_{1j}(t, \psi) \\ \vdots \\ f_{mj}(t, \psi) \end{bmatrix} + \begin{bmatrix} \frac{\partial g_1(\eta, t, \psi)}{\partial \psi_j} \\ \vdots \\ \frac{\partial g_m(\eta, t, \psi)}{\partial \psi_j} \end{bmatrix}. \quad (\text{A5})$$

These have to be solved together with equations (A1), all for some *a priori* given values of the parameter vector ψ .

We have already referenced the use by Valko and Vajda (1984) of the direct method of finding the sensitivities in chemical kinetic problems. More recently Uciński (1999) compared this method with two others, a finite-difference method and an adjoint method. He finds the direct method to be most appropriate for a wide range of applications in engineering, chemistry, physics and other experimental sciences.

The above general setup simplifies when applied to a single response function. However, it becomes more complicated when the functions η_i depend on more than one explanatory variable. Uciński (1999) presents the method for a single response function depending on time and two other factors. His scientific problems include the location of seismic sensors and pollution monitors.

A.2 Sensitivity Equations for the Reaction $A \rightarrow B \rightarrow C$

We now consider the numerical example used in this paper. The vector of the expected multivariate response is given by a solution of (3) where the concentration of each of the chemical compounds $[A]$, $[B]$ and $[C]$ is a function of t and ψ . The first equation can be solved analytically and the solution (5) substituted in the second equation. This gives the following differential equation for the concentration of the compound B

$$\frac{d[B]}{dt} = \theta_1 \{1 - (1 - \lambda_1)\theta_1 t\}^{\frac{\lambda_1}{1-\lambda_1}} - \theta_2 [B]^{\lambda_2} \quad (\text{A6})$$

This is a single differential equation ($m = 1$) and the system (A5) simplifies to one equation for each parameter. There are four parameters, i.e.,

$$\psi = (\theta_1, \theta_2, \lambda_1, \lambda_2)^T.$$

and we need to solve (A6) together with each of the following

$$\begin{aligned} \frac{df_1}{dt} &= -\theta_2 \lambda_2 [B]^{\lambda_2-1} f_1 + (1 - \theta_1 t) \{1 - (1 - \lambda_1)\theta_1 t\}^{\frac{\lambda_1}{1-\lambda_1}-1} \\ \frac{df_2}{dt} &= -\theta_2 \lambda_2 [B]^{\lambda_2-1} f_2 - [B]^{\lambda_2} \\ \frac{df_3}{dt} &= -\theta_2 \lambda_2 [B]^{\lambda_2-1} f_3 + \frac{\theta_1}{(1 - \lambda_1)^2} \{1 - (1 - \lambda_1)\theta_1 t\}^{\frac{\lambda_1}{1-\lambda_1}} \log\{1 - (1 - \lambda_1)\theta_1 t\} \\ &+ \frac{\lambda_1 \theta_1^2 t}{1 - \lambda_1} \{1 - (1 - \lambda_1)\theta_1 t\}^{\frac{\lambda_1}{1-\lambda_1}-1} \end{aligned}$$

$$\frac{df_4}{dt} = -\theta_2 \lambda_2 [B]^{\lambda_2 - 1} f_4(t, \psi) - \theta_2 [B]^{\lambda_2} \log[B]. \quad (\text{A7})$$

Equation (A6) together with the sensitivity expressions forms four systems of two equations to be solved numerically. We used the procedure “`dsolve`” of Maple V Release 4 to solve the systems for several sets of prior parameter values. The programming involved is little more than writing out the equations in a high level language similar to those of many statistical packages.

A.3 Sensitivity Equations for the Reaction $A \rightleftharpoons B \rightarrow C$

For more complicated reaction schemes, for example involving reversible reactions, the set of simultaneous equations (A1) and (A5) may also be quickly solved using the same Maple procedure.

The vector of the expected multivariate response for (4) is given by solution of the set of kinetic equations

$$\begin{aligned} \frac{d[A]}{dt} &= -\theta_1 [A]^{\lambda_1} + \theta_3 [B]^{\lambda_3} \\ \frac{d[B]}{dt} &= \theta_1 [A]^{\lambda_1} - \theta_2 [B]^{\lambda_2} - \theta_3 [B]^{\lambda_3} \\ \frac{d[C]}{dt} &= \theta_2 [B]^{\lambda_2}. \end{aligned} \quad (\text{A8})$$

Now we have six parameters: three rate constants and three orders of reaction, so that the vector ψ is equal to

$$\psi = (\theta_1, \theta_2, \theta_3, \lambda_1, \lambda_2, \lambda_3,)^T$$

This gives six systems to be solved numerically. Again, we may use the Maple procedure “`dsolve`” to obtain the solutions. To find the sensitivities for all parameters we need to solve the whole system (A8) simultaneously with (A4) for $j = 1, \dots, 6$, where

$$\frac{df_j(t, \psi)}{dt} = \begin{bmatrix} \frac{df_{1j}}{dt} \\ \frac{df_{2j}}{dt} \\ \frac{df_{3j}}{dt} \end{bmatrix},$$

the matrix $\partial g(\eta, t, \psi) / \partial \eta$ does not depend on j and is the same for all parameters,

$$\frac{\partial g(\eta, t, \psi)}{\partial \eta} = \begin{bmatrix} -\lambda_1 \theta_1 [A]^{\lambda_1 - 1} & \lambda_3 \theta_3 [B]^{\lambda_3 - 1} & 0 \\ \lambda_1 \theta_1 [A]^{\lambda_1 - 1} & -\lambda_2 \theta_2 [B]^{\lambda_2 - 1} - \lambda_3 \theta_3 [B]^{\lambda_3 - 1} & 0 \\ 0 & \lambda_2 \theta_2 [B]^{\lambda_2 - 1} & 0 \end{bmatrix},$$

$$f_j(t, \psi) = \begin{bmatrix} f_{1j} \\ f_{2j} \\ f_{3j} \end{bmatrix},$$

and the vectors $\partial g(\eta, t, \psi)/\partial \psi_j$ for the six parameters are

$$\frac{\partial g(\eta, t, \psi)}{\partial \theta_1} = \begin{bmatrix} -[A]^{\lambda_1} \\ [A]^{\lambda_1} \\ 0 \end{bmatrix}, \quad \frac{\partial g(\eta, t, \psi)}{\partial \theta_2} = \begin{bmatrix} 0 \\ -[B]^{\lambda_2} \\ [B]^{\lambda_2} \end{bmatrix}$$

$$\frac{\partial g(\eta, t, \psi)}{\partial \theta_3} = \begin{bmatrix} [B]^{\lambda_3} \\ -[B]^{\lambda_3} \\ 0 \end{bmatrix}, \quad \frac{\partial g(\eta, t, \psi)}{\partial \lambda_1} = \begin{bmatrix} -\theta_1 [A]^{\lambda_1} \ln[A] \\ \theta_1 [A]^{\lambda_1} \ln[A] \\ 0 \end{bmatrix}$$

$$\frac{\partial g(\eta, t, \psi)}{\partial \lambda_2} = \begin{bmatrix} 0 \\ -\theta_2 [B]^{\lambda_2} \ln[B] \\ \theta_2 [B]^{\lambda_2} \ln[B] \end{bmatrix}, \quad \frac{\partial g(\eta, t, \psi)}{\partial \lambda_3} = \begin{bmatrix} \theta_3 [B]^{\lambda_3} \ln[B] \\ -\theta_3 [B]^{\lambda_3} \ln[B] \\ 0 \end{bmatrix}$$

As a solution we get numerical values for the concentration of $[A]$, $[B]$ and $[C]$ and for 18 functions $f_{ij}, i = 1, 2, 3, j = 1, \dots, 6$. These 18 sensitivities would then be used as indicated in §3.4 to design a multiresponse experiment.

References

- Atkins, P. W. (1998). *Physical Chemistry*. Oxford: Oxford University Press.
- Atkinson, A. C. (2002). Non-constant variance and the design of experiments for chemical kinetic models. In S. Asprey and S. Macchietto (Eds.), *The Life of a Process Model*. Amsterdam: Elsevier. (In press).
- Atkinson, A. C. and B. Bogacka (1997). Compound, D- and D_s -optimum designs for determining the order of a chemical reaction. *Technometrics* 39, 347–356.
- Atkinson, A. C., B. Bogacka, and M. Bogacki (1998). D- and T-optimum designs for the kinetics of a reversible chemical reaction (with discussion). *Chemometrics and Intelligent Laboratory Systems* 43, 185–206.
- Atkinson, A. C. and A. N. Donev (1992). *Optimum Experimental Designs*. Oxford: Oxford University Press.
- Bates, D. M. and D. G. Watts (1988). *Nonlinear Regression Analysis and Its Applications*. New York: Wiley.

- Bauer, I., H. G. Bock, S. Körkel, and J. P. Schlöder (2000). Numerical methods for optimum experimental design in DAE systems. *Journal of Computational and Applied Mathematics* 120, 1–25.
- Box, G. E. P. and D. R. Cox (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society, Series B* 26, 211–246.
- Box, G. E. P. and H. L. Lucas (1959). Design of experiments in nonlinear situations. *Biometrika* 46, 77–90.
- Chaloner, K. and K. Larntz (1989). Optimal Bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference* 21, 191–208.
- Chaloner, K. and I. Verdinelli (1995). Bayesian experimental design: a review. *Statistical Science* 10, 273–304.
- Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics* 24, 586–602.
- Cook, R. D. and W. K. Wong (1994). On the equivalence between constrained and compound optimal designs. *Journal of the American Statistical Association* 89, 687–692.
- Draper, N. R. and W. G. Hunter (1966). Design of experiments for parameter estimation in multiresponse situations. *Biometrika* 53, 525 – 533.
- Hamilton, D. C. and D. G. Watts (1985). A quadratic design criterion for precise estimation in nonlinear regression models. *Technometrics* 27, 241–250.
- Horwitz, W. (1982). Evaluation of analytical methods used for regulation of food and drugs. *Analytical Chemistry* 54, 67A–76A.
- Hunter, W. G., W. J. Hill, and T. L. Henson (1969). Designing experiments for some or all of the constants in a mechanistic model. *Canadian Journal of Chemical Engineering* 47.
- Jennrich, R. I. and P. B. Bright (1976). Fitting systems of linear differential equations using computer generated exact derivatives (with discussion). *Technometrics* 18, 385–399.
- Kiefer, J. and J. Wolfowitz (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics* 12, 363–366.
- Lutchen, K. R. and G. M. Saidel (1982). Sensitivity analysis and experimental design techniques: application to nonlinear dynamic lung models. *Computers and Biomedical Research* 15, 434–454.

- Nathanson, M. H. and G. M. Saidel (1985). Multiple-objective criteria for optimal experimental design: application to ferrokinetics. *American Journal of Physiology* 248, R378–R386.
- Pilling, M. J. and P. W. Seakins (1995). *Reaction Kinetics*. Oxford: Oxford University Press.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes in Fortran (2nd edition)*. Cambridge, England: Cambridge University Press.
- Rocke, D. M. and S. Lorenzato (1995). A two-component model for measurement error in analytical chemistry. *Technometrics* 37, 176–184.
- Seber, G. A. F. and C. J. Wild (1989). *Nonlinear Regression*. New York: Wiley.
- Silvey, S. D. (1980). *Optimum Design*. London: Chapman and Hall.
- Uciński, D. (1999). *Measurement Optimization for Parameter Estimation in Distributed Systems*. Zielona Góra: Technical University Press.
- Valko, P. and S. Vajda (1984). An extended ODE solver for sensitivity calculations. *Computers and Chemistry* 8, 255–271.
- Wynn, H. P. (1970). The sequential generation of D -optimal experimental designs. *Annals of Mathematical Statistics* 41, 1055–1064.
- Xu, Q.-S., Y.-Z. Liang, and K.-T. Fang (2000). The effects of different experimental designs on parameter estimation in the kinetics of a reversible chemical reaction. *Chemometrics and Intelligent Laboratory Systems* 52, 155–166.

General Consecutive Reaction

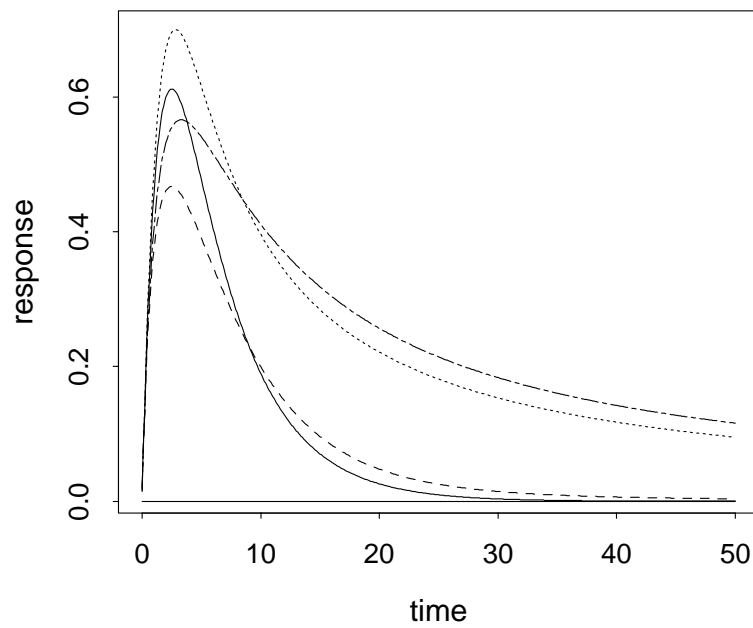


Figure 1: General consecutive reaction: response as a function of time. Reading upward at $t = 20$ $(\lambda_1, \lambda_2) = (1,1), (2,1), (1,2)$ and $(2,2)$; $\theta = (0.7, 0.2)$

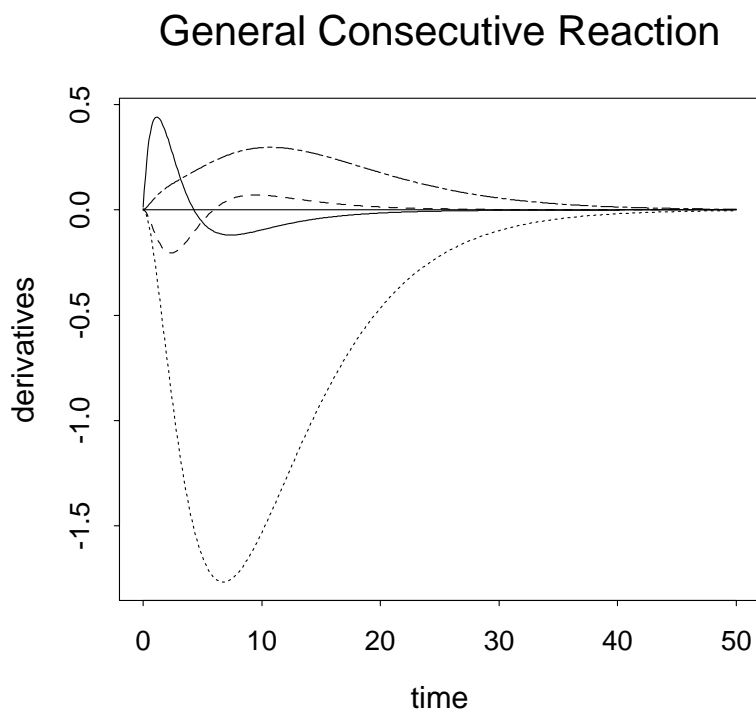


Figure 2: Derivatives (parameter sensitivities) as a function of time. Reading upward at $t = 10$: f_2 and f_1 , sensitivities for θ_2 and θ_1 ; f_3 and f_4 , sensitivities for λ_1 and λ_2

A -> B -> C: lambda = (1,1)

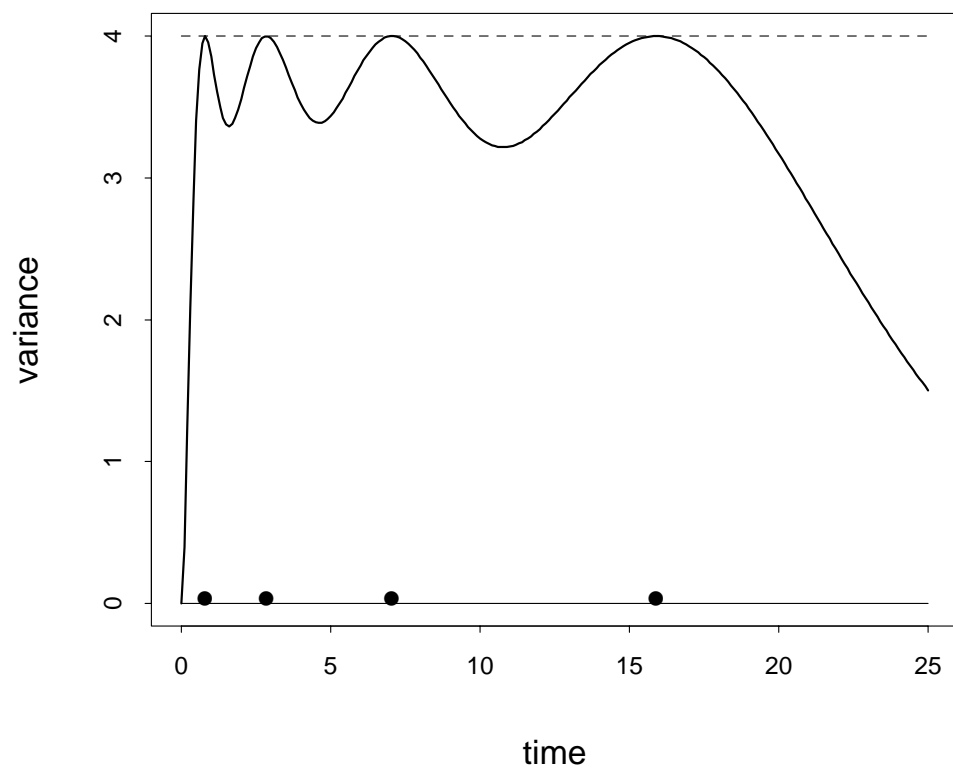


Figure 3: Variance $d(t, \xi^*, \psi)$ for the D-optimum design for rate and order when $\lambda_1^o = \lambda_2^o = 1$. The design satisfies the equivalence theorem since the maxima of the variance are at the design points and equal four

A -> B -> C: lambda = (2,2)

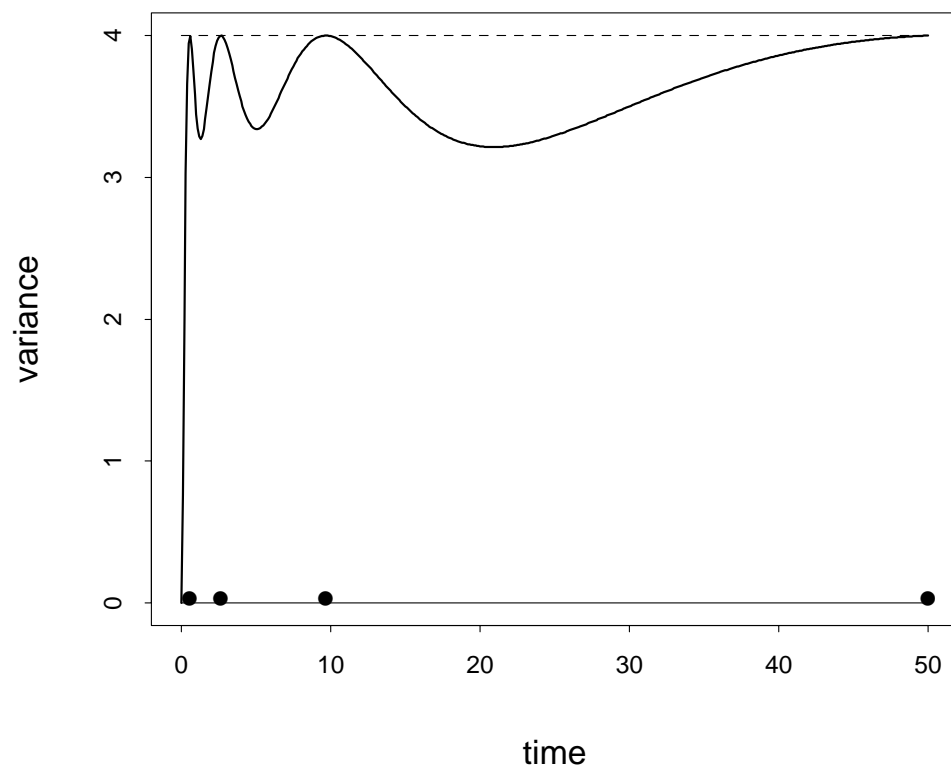


Figure 4: Variance $d(t, \xi^*, \psi)$ for the D-optimum design for rate and order when $\lambda_1^o = \lambda_2^o = 2$. The design also satisfies the equivalence theorem, but the maximum at $t = 50$ indicates that measurements could with advantage be taken at a larger time

Compound Design

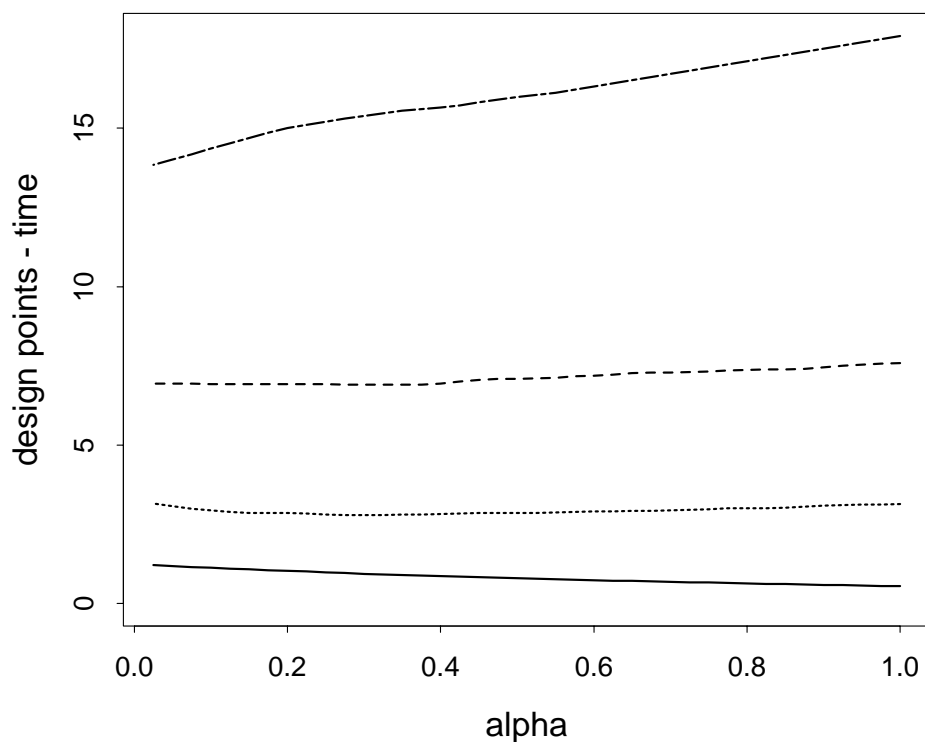


Figure 5: Support points $t_1 - t_4$ of the compound optimum design for $\lambda_1^o = \lambda_2^o = 1$ as a function of α : $\alpha = 1$ gives the D_s -optimum design for order of reaction when the rates are considered nuisance parameters

Compound Design

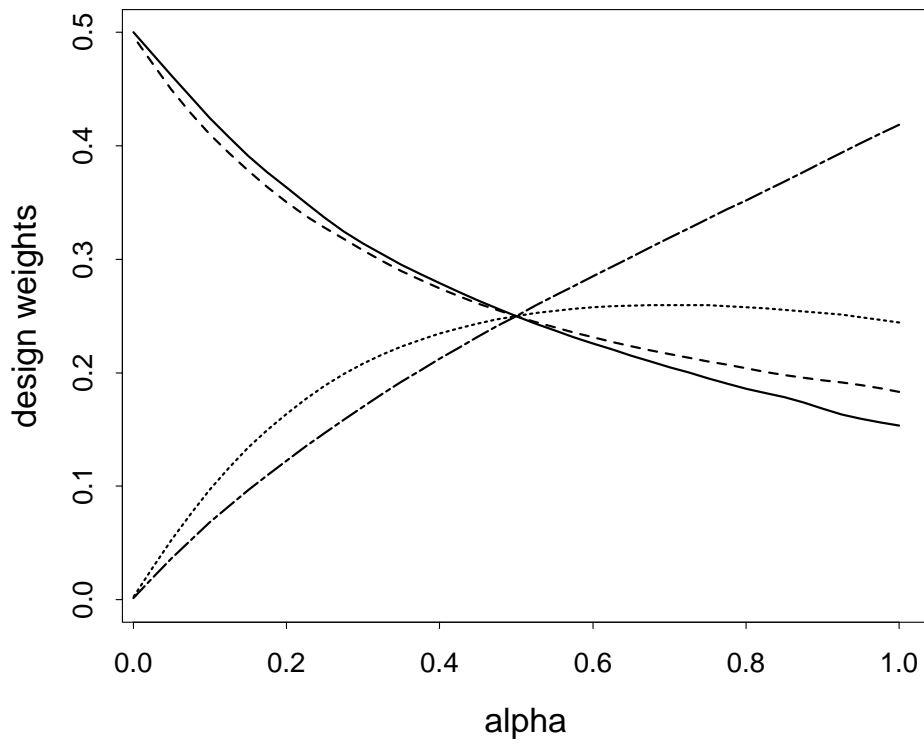


Figure 6: Design weights for the compound optimum design for $\lambda_1^o = \lambda_2^o = 1$ as a function of α . Reading upward at $\alpha = 1$: w_1, w_3, w_2 and w_4

Efficiencies of Compound Design

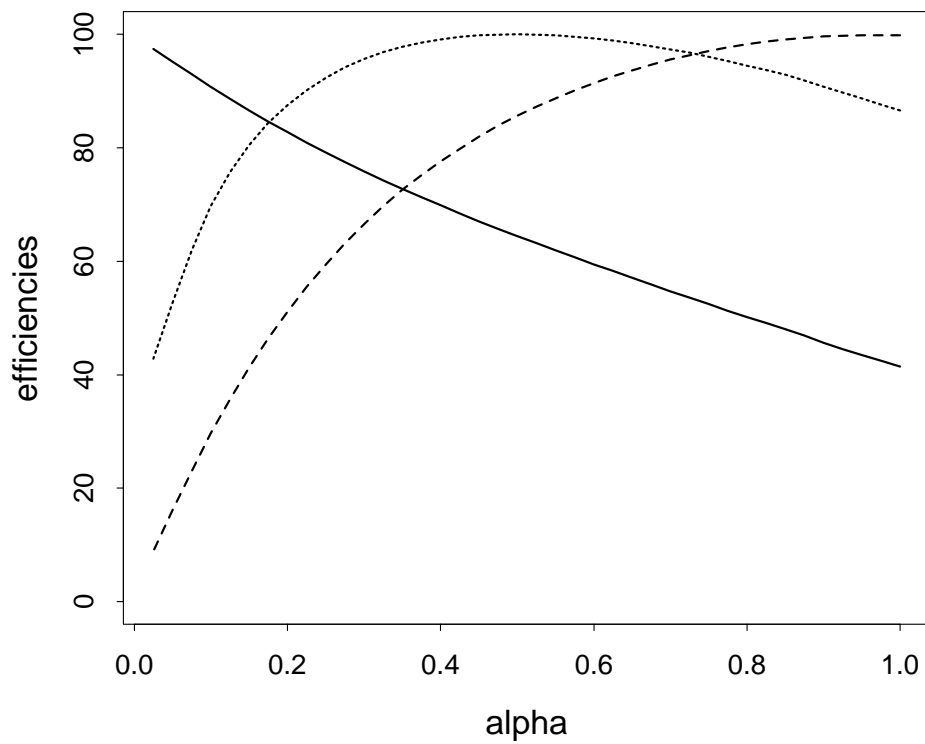


Figure 7: Efficiencies for the compound optimum design for $\lambda_1^o = \lambda_2^o = 1$ as a function of α . Reading upward at $\alpha = 1$: E_θ , E_D and E_λ

Four Lambdas, Five Design Points

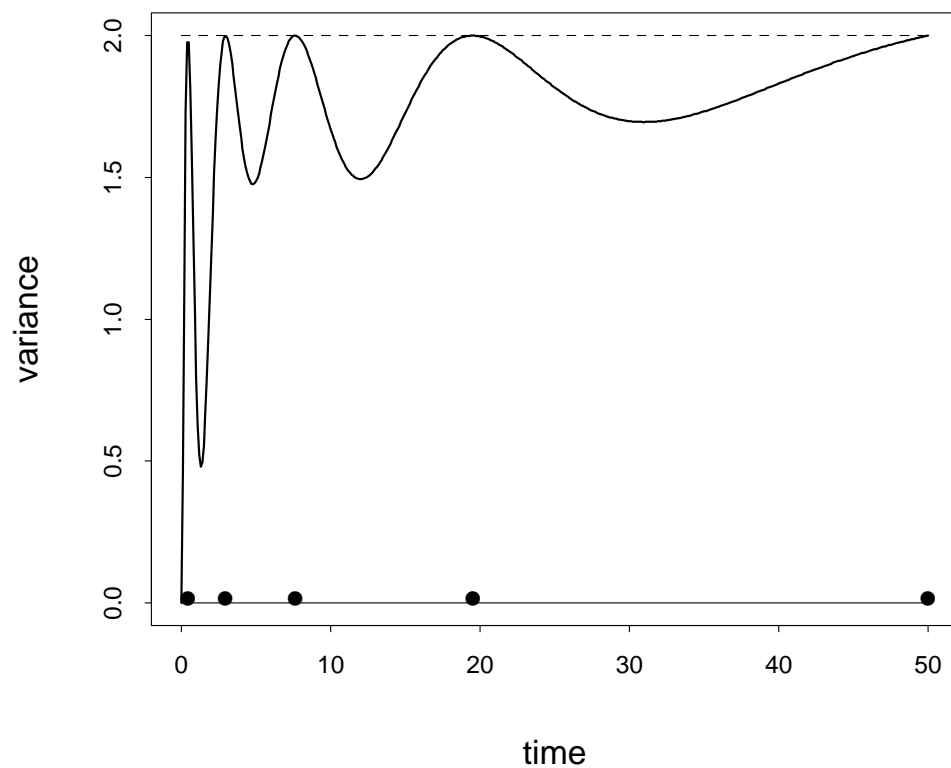


Figure 8: Variance $d(t, \xi_B^*, \psi^o, \lambda)$ for the Bayesian D_s -optimum design for order, with rates nuisance parameters, taken over four pairs of λ values. The five-point design satisfies the equivalence theorem, although a larger value of the design criterion would be obtained by extending the design region

Four Lambdas, Four Design Points

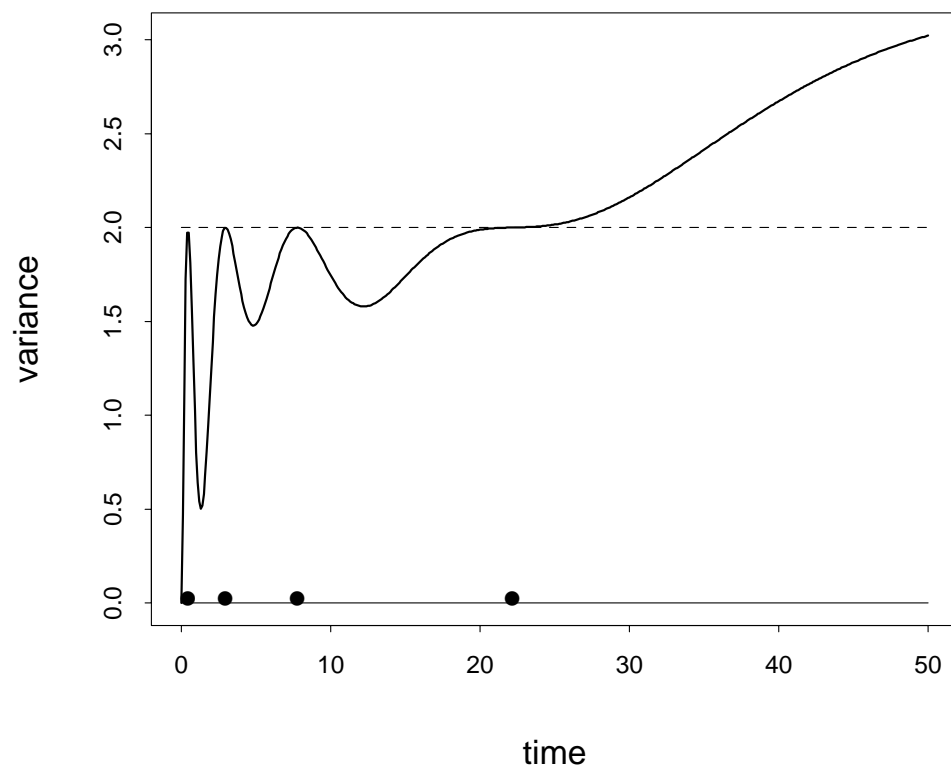


Figure 9: Variance $d(t, \xi, \psi^o, \lambda)$ for the best four-point design for order with rate a nuisance parameter taken over four pairs of λ values. The four-point design does not satisfy the equivalence theorem, although it is the best we could find from maximizing the Bayesian D_s -optimum design criterion