COMMUNICATIONS IN STATISTICS
Theory and Methods
Vol. 33, No. 5, pp. 1187–1204, 2004

Bayesian Minimally Supported *D*-Optimal Designs for an Exponential Regression Model

Zhide Fang[1,*] and Douglas P. Wiens[2]

[1]Department of Mathematics, University of New Orleans, New Orleans, Louisiana, USA
[2]Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

We consider the problem of obtaining static (i.e., nonsequential), approximate optimal designs for a nonlinear regression model with response $E[Y|x] = \exp(\theta_0 + \theta_1 x + \cdots + \theta_k x^k)$. The problem can be transformed to the design problem for a heteroscedastic polynomial regression model, where the variance function is of an exponential form with unknown parameters. Under the assumption that sufficient prior information about these parameters is available, minimally supported Bayesian *D*-optimal designs are obtained. A general procedure for constructing such designs is provided; as well the analytic forms of these designs are derived for some special

*Correspondence: Zhide Fang, Department of Mathematics, University of New Orleans, Room 207, MATH Building, 2000 Lakeshore Drive, New Orleans, LA 70148, USA; Fax: (504)-280-5516; E-mail: zfang@math.uno.edu.

DOI: 10.1081/STA-120029833 0361-0926 (Print); 1532-415X (Online)

priors. The theory of canonical moments and the theory of continued fractions are applied for these purposes.

Key Words: Canonical moments; Continued fractions; Nonlinear regression.

1. INTRODUCTION

1.1 A design problem for a nonlinear regression model. Consider the design problem for a nonlinear regression model, in which one observes, with additive, independent, homoscedastic errors, a response variable $y(x)$ with expected value

$$g(x, \theta) = \exp(\theta_0 + \theta_1 x + \cdots + \theta_k x^k), \tag{1}$$

corresponding to an input variable lying in an interval $[a, b]$, $-\infty \leq a < b \leq \infty$. The parameters θ_j are unknown, and are to be estimated.

We construct optimal designs for model (1), restricting to *approximate* designs, that is, discrete probability measures on the design space. Methods of implementing such approximate designs are discussed in Pukelsheim (1993, Chapter 12). Given a statistical model, the loss for choosing the optimal design is usually defined in terms of some scalar-valued function of the information matrix. Then the optimal design is the one which optimizes such a criterion among a class of candidate designs. When the model is nonlinear, the design problem becomes much harder than that for a linear model because the information matrix depends on the unknown parameters. In the case that sufficient prior information is available, one common way to handle this difficulty is to assume a prior distribution on the parameters, thus to seek Bayesian optimal designs.

Let ξ be any design measure. For the nonlinear model (1) the information matrix of ξ, whose inverse is proportional to the asymptotic covariance matrix of the parameter estimates, is given by

$$\int_a^b \left(\frac{\partial g(x, \theta)}{\partial \theta}\right)\left(\frac{\partial g(x, \theta)}{\partial \theta}\right)^T d\xi.$$

With $\mathbf{f}_k(x) = (1, x, \ldots, x^k)^T$, this information matrix is the same as that of a heteroscedastic kth degree polynomial regression model with efficiency function $\exp(2\mathbf{f}_k^T(x)\theta)$; this is since $\partial g(x,\theta)/\partial\theta = \exp(\mathbf{f}_k^T(x)\theta)\mathbf{f}_k(x)$. Thus, designs for (1) can be obtained from those for a heteroscedastic polynomial regression model.

1.2 A design problem for a heteroscedastic polynomial model. In this paper, more generally, we assume the polynomial regression model with

$$E[y(x)] = \sum_{j=0}^{n} \beta_j x^j, \tag{2}$$

where $x \in S = [a, b]$, $-\infty \le a < b \le \infty$, and $\mathrm{var}(y(x)) = \lambda_k^{-1}(x,\theta)$ for the efficiency function $\lambda_k(x,\theta) = \exp(\mathbf{f}_k^T(x)\theta)$. The information matrix of any design ξ for the model (2) is given by

$$M(\xi,\theta) = \int_S \mathbf{f}_n(x)\mathbf{f}_n^T(x)\lambda_k(x,\theta)\xi(dx),$$

which depends on the unknown parameter vector θ. When the efficiency function is independent of x the information matrix is proportional to

$$M_0 = \int_S \mathbf{f}_n(x)\mathbf{f}_n^T(x)\xi(dx).$$

Given any prior distribution $\pi(\theta)$ of θ on the parameter space $\Theta = \mathscr{R}^{k+1}$, a Bayesian optimal design is the one maximizing

$$\Psi(\xi) = \int_\Theta \log|M(\xi,\theta)|\, d\pi(\theta) \tag{3}$$

among all designs ξ. If $\pi(\theta)$ places unit mass on some known vector θ^*, this reduces to the D-optimality criterion for weighted polynomial regression with a specified weight function. See Chaloner and Verdinelli (1995), Chaloner and Clyde (1996) for further discussion about the motivation of this criterion.

Bayesian optimal designs for different nonlinear models (for example, exponential growth models) are addressed in Chaloner (1993), Mukhopadhyay and Haines (1995), Dette and Neugebauer (1996, 1997), Dette and Wong (1996, 1998) and others. Alternative methods to construct optimal designs for nonlinear models are also available in the literature. One of them is to adopt a maximin approach, which maximizes, over designs ξ, the minimum of the determinant of the information matrix over $\theta \in \Theta$. Imhof (2001) obtains maximin D-optimal designs for different models. See Sinha and Wiens (2002) for robustness aspects of this problem.

2. PRELIMINARIES

We briefly review the concepts and some basic properties of the canonical moments of a probability measure. These will be utilized in Sec. 3 for constructing Bayesian optimal designs. We adopt notation as in Dette and Studden (1997), henceforth referred to as DS. For an arbitrary probability measure ξ with support S, let $c_j = \int_S x^j \xi(dx)$, $j = 0, 1, \ldots$, be the ordinary moments of ξ. Denote by c_j^+ (c_j^-) the maximum (minimum) value of c_j, given $c_0, \ldots, c_{j-1}$. Then the canonical moments are defined as $p_j = (c_j - c_j^-)/(c_j^+ - c_j^-)$, $j = 1, 2, \ldots$, whenever $c_j^- < c_j^+$. The correspondence between sequences $\{c_j\}_{j \geq 1}$ and $\{p_j\}_{j \geq 1}$ is one-to-one. A probability measure ξ has finite support if and only if the corresponding sequence of canonical moments terminates, that is, $p_n = 0$ or 1 for some integer n.

For any ξ, with canonical moments $p_1, p_2, \ldots$, set $q_j = 1 - p_j$ and $\zeta_0 = 1$, $\zeta_1 = p_1$, $\zeta_j = q_{j-1} p_j$, $j \geq 2$. Then (DS, p. 149)

$$|M_0(\xi)| = (b-a)^{n(n+1)} \prod_{j=1}^{n} (\zeta_{2j-1}\zeta_{2j})^{n-j+1}. \tag{4}$$

The Stieltjes transform of ξ on $S = [a, b]$ has the continued fraction expansion (DS, p. 89)

$$\begin{aligned}\int_S \frac{\xi(dx)}{z-x} &= \frac{1\ \ |}{|z-a} - \frac{\zeta_1(b-a)|}{|\ \ \ 1} - \frac{\zeta_2(b-a)|}{|\ \ z-a} - \cdots \\ &= \frac{1}{z-a-\frac{\zeta_1(b-a)}{1-\frac{\zeta_2(b-a)}{z-a-\cdots}}},\end{aligned} \tag{5}$$

for any $z \in \mathscr{C} \backslash S$, where $\mathscr{C}$ denotes the complex plane. If ξ has finite support, with support points $\{x_1, \ldots, x_{n+1}\}$, then $\zeta_{2n+1}\zeta_{2n} = 0$ (see for example Theorem 2.2.3 of DS). Then the above expansion has at most $2n+2$ terms and can be expressed as the ratio of two polynomials. The support points of ξ coincide with the zeros of the polynomial in the denominator of this ratio. Furthermore, if $D_{n+1}(z)$ is the polynomial in the denominator, then we can write

$$D_{n+1}(z) = \begin{vmatrix} z-a & -1 & 0 & \cdots & 0 & 0 \\ -\zeta_1(b-a) & 1 & -1 & \cdots & 0 & 0 \\ 0 & -\zeta_2(b-a) & z-a & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & z-a & -1 \\ 0 & 0 & 0 & \cdots & -\zeta_{2n+1}(b-a) & 1 \end{vmatrix}$$

$$= z^{n+1} - \left[(n+1)a + \sum_{j=1}^{2n+1} \zeta_j(b-a)\right] z^n + \left[\frac{n(n+1)}{2}a^2 + na(b-a). \times \sum_{j=1}^{2n+1} \zeta_j + (b-a)^2\left(\sum_{i<j}^{2n+1} \zeta_i\zeta_j - \sum_{j=1}^{2n} \zeta_j\zeta_{j+1}\right)\right] z^{n-1} + \cdots.$$

Thus, in light of Newton's identities relating the zeros and the coefficients of a polynomial (MacDuffee 1962, Theorem 48), we have the equations

$$\sum_{i=1}^{n+1} x_i = (n+1)a + \sum_{j=1}^{2n+1} \zeta_j(b-a), \tag{6}$$

$$\sum_{i=1}^{n+1} x_i^2 = (n+1)a^2 + 2a(b-a)\sum_{j=1}^{2n+1} \zeta_j + (b-a)^2\left(\sum_{j=1}^{2n+1} \zeta_j^2 + 2\sum_{j=1}^{2n} \zeta_j\zeta_{j+1}\right). \tag{7}$$

3. BAYESIAN *D*-OPTIMAL DESIGNS

Assume the model (2) and a prior distribution $\pi(\theta)$ on θ. We seek the Bayesian optimal design, which maximizes the criterion $\Psi(\xi)$ given in (3). In order that the regression parameters $\{\beta_0, \beta_1, \ldots, \beta_n\}$ be estimable, a design ξ must have at least $n+1$ support points. To reduce the algebraic complexity, while maintaining adequate generality for most applications, we assume $k = 2$.

There is no restriction on the number of observations to be collected. However, the number of levels of x will be restricted to $n+1$. This is motivated by the following property, which as pointed out by Dette and Wong (1998) can be established by an argument similar to that in Theorem 5.1, Karlin and Studden (1966).

Property 1. If the prior $\pi(\theta)$ places unit mass at $\theta^* \in \Theta$, then the Bayesian D-optimal design has $n+1$ support points.

Denote by Ξ_{n+1} the class of all designs with $n+1$ support points. In this paper, we construct Bayesian D-optimal designs within Ξ_{n+1} for any given $\pi(\theta)$.

For any $\xi \in \Xi_{n+1}$, with design points $\{x_1, \ldots, x_{n+1}\}$, let F be the square matrix with ith row being $(1, x_i, \ldots, x_i^n)$, $i = 1, 2, \ldots, n+1$. Then, by Corollary 1 in Fedorov (1972, p. 84), we have

$$|M(\xi, \theta)| = \left\{\prod_{i=1}^{n+1} \xi(x_i)\right\}\left\{\prod_{i=1}^{n+1} \lambda_2(x_i, \theta)\right\}|F|^2. \tag{8}$$

From this it follows that the Bayesian D-optimal design within Ξ_{n+1} concentrates equal mass on each of its support points, that is, $\xi(x_i) = 1/(n+1)$ for $i = 1, 2, \ldots, n+1$.

The following theorem shows that the Bayesian D-optimal design within Ξ_{n+1} depends only on the first marginal moments of $\pi(\theta)$. Denote these by $E_i = \int_\Theta \theta_i \, d\pi(\theta)$ for $i = 0, 1, 2$.

Theorem 1. *Assume the model* (2) *and any prior* $\pi(\theta)$. *Let* $q_{-1} = 0$ *and* $q_0 = 1$.

(I) *If* $E_2 < 0$ *and* $E_1 + 2aE_2 \leq 0$ *then the Bayesian D-optimal design within* Ξ_{n+1} *has canonical moment* $p^*_{2n+1} = 0$, *and*

(CM1) *Canonical moments* $p^*_1, \ldots, p^*_{2n} \in (0, 1)$ *determined by the* $2n$ *equations*

$$\frac{n-j+1}{p_{2j-1}q_{2j-2}} = -(b-a)(E_1 + 2aE_2) - 2(b-a)^2 \times E_2[q_{2j-3}p_{2j-2} + q_{2j-2}p_{2j-1} + q_{2j-1}p_{2j}], \quad j = n, \ldots, 1, \tag{9}$$

$$\frac{n-j+1}{p_{2j}q_{2j-1}} = -(b-a)(E_1+2aE_2) - 2(b-a)^2 \times E_2[q_{2j-2}p_{2j-1} + q_{2j-1}p_{2j} + q_{2j}p_{2j+1}], \quad j = n, \ldots, 1, \tag{10}$$

if the system has a solution in the interior of $[0,1]^{2n}$. *Otherwise,*

(CM2) *Canonical moments* $p_1^*, \ldots, p_{2n-1}^* \in (0,1)$, $p_{2n}^* = 1$ *determined by the* $2n-1$ *equations*

$$\frac{n-j+1}{p_{2j-1}} - \frac{n-j+1}{q_{2j-1}} = -(b-a)(E_1+2aE_2)(q_{2j-2}-p_{2j}) - 2(b-a)^2 E_2[q_{2j-2}^2 p_{2j-1} - q_{2j-1}p_{2j}^2 + q_{2j-3}p_{2j-2}q_{2j-2} + q_{2j-2}(q_{2j-1}-p_{2j-1})p_{2j} - p_{2j}q_{2j}p_{2j+1}], \quad j = n, n-1, \ldots, 1, \tag{11}$$

$$\frac{n-j+1}{p_{2j}} - \frac{n-j}{q_{2j}} = -(b-a)(E_1+2aE_2)(q_{2j-1}-p_{2j+1}) - 2(b-a)^2 E_2[q_{2j-1}^2 p_{2j} - q_{2j}p_{2j+1}^2 + q_{2j-2}p_{2j-1}q_{2j-1} + q_{2j-1}(q_{2j}-p_{2j})p_{2j+1} - p_{2j+1}q_{2j+1}p_{2j+2}], \quad j = n-1, \ldots, 1. \tag{12}$$

(II) *If* $E_2 \geq 0$ *and* $E_1 + 2aE_2 > 0$, *then the Bayesian D-optimal design within* Ξ_{n+1} *has canonical moments* $p_{2n+1}^* = 1$, *and*

(CM3) *Canonical moments* $p_1^*, \ldots, p_{2n}^* \in (0,1)$ *determined by the* $2n$ *equations*

$$\frac{n-j+1}{p_{2j}p_{2j-1}} = (b-a)(E_1+2aE_2) + 2(b-a)^2 \times E_2[p_{2j-2}q_{2j-1} + p_{2j-1}q_{2j} - q_{2j}q_{2j+1}], \quad j = n, \ldots, 2, 1, \tag{13}$$

$$\frac{n-j+1}{q_{2j-1}q_{2j-2}} = (b-a)(E_1+2aE_2) + 2(b-a)^2$$
$$E_2[p_{2j-2}q_{2j-1} + p_{2j-1}q_{2j} - p_{2j-3}p_{2j-2}],$$
$$j = n, \ldots, 2, 1, \tag{14}$$

if the system has a solution in the interior of $[0,1]^{2n}$. *Otherwise,*

(CM4) *Canonical moments* $p_1^*, \ldots, p_{2n-1}^* \in (0,1)$, $p_{2n}^* = 1$, *determined by* (11), (12).

Proof. From Eqs. (4), (8), if $\xi \in \Xi_{n+1}$, we have

$$|M(\xi,\theta)| = \left(\prod_{i=1}^{n+1} \lambda_2(x_i,\theta)\right)|M_0(\xi)|$$
$$= (b-a)^{n(n+1)} \exp\left((n+1)\theta_0 + \theta_1 \sum_{i=1}^{n+1} x_i + \theta_2 \sum_{i=1}^{n+1} x_i^2\right)$$
$$\times \prod_{j=1}^{n} (\zeta_{2j-1}\zeta_{2j})^{n-j+1}.$$

Then, in light of (6) and (7),

$$\Psi(\xi) = d + \sum_{j=1}^{n}(n-j+1)\log(q_{2j-2}p_{2j-1}q_{2j-1}p_{2j})$$
$$+ (b-a)(E_1+2aE_2)\sum_{j=1}^{2n+1} q_{j-1}p_j$$
$$+ (b-a)^2 E_2\left[\sum_{j=1}^{2n+1}(q_{j-1}p_j)^2 + 2\sum_{j=1}^{2n}(q_{j-1}p_jq_jp_{j+1})\right]$$

is a function of $(p_1, \ldots, p_{2n+1})$, where $d = (n+1)(n\log(b-a) + E_0 + aE_1 + a^2E_2)$ does not depend on ξ. To see the existence of the Bayesian D-optimal design within Ξ_{n+1}, consider the problem of maximizing $\Psi(\xi)$ over $(p_1, \ldots, p_{2n+1})$. Since the objective function is a continuous function on the compact space $[0,1]^{2n+1}$, there are solutions to this problem. Let $(p_1^*, \ldots, p_{2n}^*, p_{2n+1}^*)$ be one solution. It is obvious that $0 < p_i^* < 1$ for $i = 1, \ldots, 2n-1$ and $p_{2n}^* \neq 0$ since otherwise

$\Psi(\xi^*) = -\infty$. So the design ξ^* corresponding to $(p_1^*, \ldots, p_{2n}^*, p_{2n+1}^*, 0)$ must have $n+1$ support points.

(I) When $E_2 < 0$ and $E_1 + 2aE_2 \leq 0$, it is obvious that $p_{2n+1}^* = 0$. Assume $p_{2n}^* < 1$. By differentiating the objective function with respect to $p_i, i = 1, \ldots, 2n$, we have that $(p_1^*, \ldots, p_{2n}^*)$ is the solution set to the system consisting of the equations

$$\frac{1}{p_{2n}q_{2n-1}} = -(b-a)(E_1 + aE_2) - 2(b-a)^2 E_2[q_{2n-2}p_{2n-1} + q_{2n-1}p_{2n}], \tag{15}$$

and (11), (12). Similar to the arguments in Dette and Wong (1998), we obtain the system stated in (CM1) by sequentially inserting (15) into (11), the resulting equation into (12), and so on.

If $p_{2n}^* = 1$, then by differentiating the objective function with respect to $p_1, \ldots, p_{2n-1}$, we obtain the system stated in (CM2). So $(p_1^*, \ldots, p_{n-1}^*)$ is the solution set to this system.

(II) It is easy to see that $p_{2n+1}^* = 1$ under the assumed conditions. Equations (13) and (14) are obtained by arguments similar to those used to establish (CM1). □

Remarks. 1. From the examples presented next, we observe that for the case in Theorem 1(I), given E_2, there exist a function $\gamma(E_2)$, such that if $E_1 < \gamma_1(E_2)$ then the canonical moments of the Bayesian D-optimal design within Ξ_{n+1} can be found by solving the system in (CM1) of Theorem 1. Otherwise, they are solutions to the system in (CM2). In fact, from the last equation in (CM1), one can see that $p_{2n}^* < 1$ is equivalent to

$$E_1 < -\left(\frac{1}{q_{2n-1}^*(b-a)} + 2E_2\left[(b-a)(q_{2n-1}^* + q_{2n-2}^* p_{2n-1}^*) + a\right]\right),$$

which partially verifies the observation. The explicit expression for $\gamma(E_2)$ remains outstanding except for the simplest case when $n = 1$ (see Example 1). Similarly, for the case in Theorem 1(II), given E_2, there exists a function $\gamma_2(E_2)$, such that if $E_1 > \gamma_2(E_2)$ then the canonical moments of the Bayesian D-optimal design within Ξ_{n+1} can be found by solving the system in (CM3) of Theorem 1. Otherwise, they are solutions to the system in (CM2).

2. The above theorem considers only the cases when $E_1 + aE_2$ and E_2 have the same sign. If $E_1 + aE_2 \leq 0$ and $E_2 > 0$, then $p_{2n+1}^* = 0$ or 1.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Other canonical moments can be found by (I) if $p_{2n+1}^* = 0$ or (II) if $p_{2n+1}^* = 1$. When $E_1 + aE_2 \geq 0$ and $E_2 < 0$, there exists the case that $p_{2n+1}^* \in (0,1)$. Note that if $p_{2n}^*, p_{2n+1}^* \in (0,1)$, the design ξ^* corresponding to canonical moments $(p_1^*, \ldots, p_{2n+1}^*, 0)$ is minimally supported. Furthermore, these canonical moments can be obtained by (9), (10) and

$$(E_1 + 2aE_2) + 2(b-a)E_2[q_{2n}p_{2n+1} + q_{2n-1}p_{2n}] = 0.$$

3. It is easy to solve the equations in the theorem by using the built-in functions *NSolve* or *FindRoot* in *Mathematica*. With terminating sequences of canonical moments of a design, there are standard methods to locate the support points of the design. See, for example, Theorem 3.6.1 in DS, which states that if a design ξ has $n+1$ support points and canonical moments p_j and ζ_j, $j \geq 1$, then these support points are zeros of a polynomial $W_{n+1}(x)$, defined by the recursive formula

$$\begin{aligned} W_{j+1}(x) &= (x - a - (b-a)(\zeta_{2j} + \zeta_{2j+1})) \\ &\quad \times W_j(x) - (b-a)^2\zeta_{2j-1}\zeta_{2j}W_{j-1}, \quad j = 0,1,2,\ldots, \end{aligned} \tag{16}$$

with $\zeta_0 = 0$ and $W_{-1}(x) \equiv 0$, $W_0(x) \equiv 1$.

Example 1. Simple linear heteroscedastic regression models. Let $n = 1$. We assume that $E_1 + 2aE_2 \leq 0$ and $E_2 \leq 0$ (other cases can be addressed in the same manner). Then $p_3^* = 0$. Simple algebraic operations show that if $E_1 < -(2/(b-a) + 2bE_2)$, we have (from Theorem 1(I))

$$p_1^* = \begin{cases} \frac{-(E_1+2aE_2)-\sqrt{(E_1+2aE_2)^2-16E_2}}{8(b-a)E_2}, & \text{when } E_2 < 0, \\ -\frac{1}{(b-a)E_1}, & \text{when } E_2 = 0, \end{cases}$$

$$p_2^* = \frac{p_1^*}{1-p_1^*}.$$

Otherwise, we have (from Theorem 1(II))

$$p_1^* = \frac{1}{2}, \quad p_2^* = 1.$$

Thus, we have $\zeta_1^* = p_1^* = \zeta_2^*$, $\zeta_3^* = 0$. So (16) gives

$$\begin{aligned} W_2(x) &= (x - a - (b-a)p_1^*)^2 - (b-a)^2(p_1^*)^2 \\ &= (x-a)(x - a - 2(b-a)p_1^*). \end{aligned}$$

This implies that the Bayesian D-optimal design within Ξ_2 for the simple linear heteroscedastic regression model puts equal mass on each of two points

$$x_1^* = a, \quad x_2^* = a + 2(b-a)p_1^*.$$

One can see that when $p_1^* = 1/2$, then $x_1^* = a$, $x_2^* = b$. Thus, if the prior has first moments E_1, E_2 such that $-(2/(b-a) + 2bE_2) \le E_1 < -2bE_2$ and $E_2 \le 0$, the Bayesian D-optimal design within Ξ_2 for the linear heteroscedastic regression model puts equal mass on each of the two end-points of the design interval, which concurs with the D-optimal design for the simple linear regression model.

Example 2. Quadratic heteroscedastic regression model. Let $n = 2$. For simplicity, we take $a = 0, b = 1$. If $E_1 < 0, E_2 < 0$, then $p_5^* = 0$. The top part of Table 1 lists the canonical moments and corresponding design points of Bayesian D-optimal designs within Ξ_3 for some negative E_1, E_2. We observe that given E_2, if E_1 is small, the vector $(p_1^*, p_2^*, p_3^*, p_4^*)$

Table 1. Some numerical examples of the Bayesian D-optimal designs within Ξ_3 for the quadratic heteroscedastic model.

E_2	E_1	$(p_1^*, p_2^*, p_3^*, p_4^*)$	(x_1^*, x_2^*, x_3^*)[a]
−0.25	−10	(0.196, 0.243, 0.130, 0.114)	(0, 0.125, 0.463)
	−4.3	(0.424, 0.719, 0.754, 0.902)	(0, 0.277, 0.995)
	−4.2	(0.427, 0.725, 0.775, 1)	(0, 0.281, 1)
	−0.05	(0.494, 0.667, 0.524, 1)	(0, 0.482, 1)
−1	−10	(0.186, 0.225, 0.120, 0.110)	(0, 0.121, 0.438)
	−3	(0.430, 0.692, 0.688, 0.814)	(0, 0.305, 0.985)
	−2.9	(0.437, 0.710, 0.740, 1)	(0, 0.311, 1)
	−0.05	(0.481, 0.671, 0.577, 1)	(0, 0.442, 1)
0.25	10	(0.806, 0.242, 0.872, 0.110)	(0.540, 0.878, 1)
	4.7	(0.581, 0.736, 0.207, 0.965)	(0.001, 0.743, 1)
	4.6	(0.580, 0.736, 0.204, 1)	(0, 0.739, 1)
	0.05	(0.506, 0.667, 0.474, 1)	(0, 0.519, 1)
1	15	(0.879, 0.139, 0.930, 0.064)	(0.713, 0.924, 1)
	4.6	(0.591, 0.754, 0.181, 0.961)	(0.002, 0.771, 1)
	4.5	(0.590, 0.755, 0.173, 1)	(0, 0.769, 1)
	0.05	(0.524, 0.673, 0.403, 1)	(0, 0.573, 1)

[a]Designs place mass 1/3 at each point x_i^*.

of canonical moments is the solution of the system in (CM1) of Theorem 1 and the support points of the Bayesian D-optimal design ξ^* within Ξ_3 includes only one (left) end point of the design interval, otherwise, it is the solution of the system in (CM2) of Theorem 1 and the support points of ξ^* include the two end points.

If $E_1 > 0,\ E_2 > 0$, then $p_5^* = 1$. The lower part of Table 1 lists the canonical moments and corresponding design points of Bayesian D-optimal designs within Ξ_3 for some positive E_1, E_2. We have similar observation as in the case when $E_1 < 0, E_2 < 0$, except that the designs for large E_1 include the right end point of the interval.

For a general prior $\pi(\theta)$, Theorem 1 only enables us to find the minimally supported Bayesian D-optimal design numerically for the heteroscedastic polynomial model of degree n (>1). However, for some special priors, it is possible to obtain the analytical forms of the designs. Dette and Wong (1998) consider the special case when $E_2 = 0$ and the design space $S = [0, b]$ and prove that the Bayesian D-optimal design within Ξ_{n+1} puts equal mass at the zeros of the polynomial $xL_n^{(1)}(xE_1)$, where $L_n^{(1)}(x)$ is the Laguerre polynomial of degree n.

In the following theorem, we consider the special case when $E_1 = 0$ and $b = -a > 0$. By (8), we see that for any $\xi \in \Xi_{n+1}$, $\Psi(\xi)$ is the same as that for the model (2) with the efficiency function $\lambda_2(x, \theta) = \exp(\theta_0 + \theta_2 x^2)$. If $\tilde{\xi}(x) = \xi(-x)$ for a design ξ, then it is straightforward to show that $|M(\xi, \theta)| = |M(\tilde{\xi}, \theta)|$. Hence, the convexity of the criterion ensures that we can assume ξ to be symmetric.

Denote by $H_n(x)$ the Hermite polynomial of degree n, defined by the recursive formula

$$H_n(x) = 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x), \quad n = 1, 2, 3, \ldots,$$

with $H_{-1}(x) = 0,\ H_0(x) = 1$. See Szegö (1959) for detailed properties of these polynomials. Denote by $r_{\max}^{(n)}$ the maximum zero of $H_n(x)$.

Theorem 2. *Assume the model* (2) *with the design space* $S = [-b, b]$ *and any prior* $\pi(\theta)$ *with* $E_1 = 0$.

(M1) *If* $E_2 \le -(r_{\max}^{(n+1)}/b)^2$, *then the Bayesian D-optimal design within* Ξ_{n+1} *puts equal mass on each of the zeros of the Hermite polynomial* $H_{n+1}(\sqrt{|E_2|}x)$.

(M2) *If* $E_2 > -(r_{\max}^{(n+1)}/b)^2$, *then the Bayesian D-optimal design within* Ξ_{n+1} *has canonical moments* $p_{2j-1}^* = 1/2,\ j = 1, \ldots, n,$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

$p_{2n}^*=1$, and $p_2^*,\dots,p_{2n-2}^*\in(0,1)$ consisting of the solution set of the system

$$\frac{n-j+1}{p_{2j}}-\frac{n-j}{q_{2j}}+2b^2E_2(q_{2j-2}-p_{2j+2})=0,\quad j=1,2,\dots,n-1, \tag{17}$$

with $q_0=1$.

Proof. For any symmetric design $\xi\in\Xi_{n+1}$, with design points $\{x_1,\dots,x_{n+1}\}$ and canonical moments $\{p_1,\dots,p_{2n+1}\}$, we have $p_{2j-1}=1/2$ (Corollary 1.3.4 of DS), thus $\zeta_{2j}+\zeta_{2j+1}=1/2$, $j=1,\dots,n$. This, (6)—note that $\sum_{j=1}^{n+1}x_j=0$—and (7) imply that

$$\begin{aligned}\sum_{j=1}^{n+1}x_j^2 &= (n+1)b^2+2(-b)(n+1)b\\ &\quad+4b^2\left[\zeta_1^2+\sum_{j=1}^{n}(\zeta_{2j}+\zeta_{2j+1})^2+2\sum_{j=1}^{n}\zeta_{2j-1}\zeta_{2j}\right]\\ &=8b^2\sum_{j=1}^{n}\zeta_{2j-1}\zeta_{2j}=2b^2\sum_{j=1}^{n}q_{2j-2}p_{2j}.\end{aligned}$$

Therefore, the objective function of the maximization problem becomes

$$\Psi(\xi)=d_1+\sum_{j=1}^{n}(n-j+1)\log(q_{2j-2}p_{2j})+2b^2E_2\sum_{j=1}^{n}(q_{2j-2}p_{2j}),$$

a function of $\{p_2,\dots,p_{2n}\}$ in which $d_1=(n+1)(n\log b-E_0)$ does not depend on the design. As in the proof of Theorem 1, one finds that $\Psi(\xi)$ has a maximizer, say, $(p_2^*,\dots,p_{2n}^*)$, with $p_{2j}^*\in(0,1)$, $j=1,\dots,n-1$, and $p_{2n}^*\neq 0$. The design ξ^* corresponding to these canonical moments is the Bayesian D-optimal design within Ξ_{n+1}.

We claim that p_{2n}^* must be 1 if $0>E_2>-(r_{\max}^{(n+1)}/b)^2$ (it is obviously true when $E_2>0$). If $p_{2n}^*<1$, differentiating $\Psi(\xi)$ with respect to p_{2j}, $j=1,\dots,n$, equating these to 0 and solving the resulting system yields

$$q_{2j-2}^*p_{2j}^*=-\frac{n-j+1}{2b^2E_2},\quad j=1,\dots,n,$$

with $q_0^* = 1$. Thus, the Stieltjes transform of ξ^* has expansion (by the even contraction of the continued fraction in (5))

$$\int_{-b}^{b} \frac{\xi^*(dx)}{z-x} = \frac{1\ |}{|\ z} - \frac{\zeta_1^*\zeta_2^* 4b^2|}{|\ z} - \frac{\zeta_3^*\zeta_4^* 4b^2|}{|\ z} - \cdots - \frac{\zeta_{2n-1}^*\zeta_{2n}^* 4b^2|}{|\ z}$$
$$= \frac{1\ |}{|\ z} + \frac{n/(2E_2)|}{|\ z} + \frac{(n-1)/(2E_2)|}{|\ z} + \cdots + \frac{1/(2E_2)|}{|\ z}.$$

Then, the support points of ξ^* are the zeros of the polynomial

$$B_{n+1}(z) = \begin{vmatrix} z & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \frac{n}{2E_2} & z & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \frac{n-1}{2E_2} & z & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \frac{2}{2E_2} & z & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \frac{1}{2E_2} & z \end{vmatrix}.$$

These polynomials $\{B_n(x)\}_{n\geq 1}$ satisfy the recursive formula

$$B_{n+1}(z) = zB_n(z) + \frac{n}{2E_2}B_{n-1}(z),$$
$$= zB_n(z) - \frac{n}{2|E_2|}B_{n-1}(z), \quad n = 1, 2, \ldots,$$

with $B_1(z) = z$, $B_0(z) = 1$. Define $H_n(\sqrt{|E_2|}z) = 2^n|E_2|^{n/2}B_n(z)$. Then, we have

$$H_{n+1}(\sqrt{|E_2|}z) = 2\sqrt{|E_2|}zH_n(\sqrt{|E_2|}z) - 2nH_{n-1}(\sqrt{|E_2|}z),$$
$$n = 1, 2, \ldots,$$

with $H_1(\sqrt{|E_2|}z) = 2\sqrt{|E_2|}z$, $H_0(\sqrt{|E_2|}z) = 1$. Thus $H_{n+1}(\sqrt{|E_2|}z)$ must be the Hermite polynomial of degree $n+1$ and its zeros concur with the support of ξ^*. Since the design space $S = [-b, b]$, we conclude that $E_2 \leq -(r_{\max}^{(n+1)}/b)^2$. This is a contradiction and so proves the claim. Thus, under the condition in (II), canonical moments $(p_2^*, \ldots, p_{2n-2}^*)$ is the solution set to the system obtained by differentiating $\Psi(\xi)_{|_{p_{2n}=1}}$ with respect to p_{2j}, $j = 1, \ldots, n-1$, which is the system in (M2). □

We present some numerical examples of Bayesian D-optimal designs within Ξ_{n+1} in Table 2. Here, we only consider the case when $E_2 < 0$

Table 2. The maximum zeros of Hermite polynomials and the support points of Bayesian D-optimal designs within Ξ_{n+1}, with $S=[-1,1]$.

n	$r_{\max}^{(n+1)}$	E_2	$(x_1^*,\ldots,x_{n+1}^*)$
1	$\sqrt{2}/2$	-1	$(-\sqrt{2}/2,\ \sqrt{2}/2)$
		-0.4^*	$(-1,\ 1)$
2	$3/2$	-3	$(-1/2,\ 0,\ 1/2)$
		-2^*	$(-1,\ 0,\ 1)$
3	$\sqrt{(3+\sqrt{6})/2}$	-3	$\pm\sqrt{(3+\sqrt{6})/6}, \pm\sqrt{(3-\sqrt{6})/6}$
		-1	$(-1, \pm\sqrt{7-\sqrt{41}}/2, 1)$
		-0.5	$(-1, \pm(3-2\sqrt{2}), 1)$

Note: The designs remain unchanged when the values of E_2 with a "$*$" change, as long as $b\sqrt{|E_2|} < r_{\max}^{(n+1)}$.

(the case when $E_2 > 0$ is very similar). When $|E_2|$ is small, the design points include ± 1 and are obtained by calculating the zeros of the polynomial in (16) (case (M2) of Theorem 2). □

4. CASE STUDY

In modelling allometric growth curves, attention focuses on applications in which the ratio between increments in, typically biological, structures remains approximately constant. A log–log transformation may then yield a linear relationship between mean logged response and logged regressor, at the cost of destroying the original additive error structure. See Griffiths and Sandland (1984) for a discussion. Prange et al. (1979), in an experiment described also in Sokal and Rohlf (1995, p. 553), study the relationship between body mass (Y) and skeletal mass (Z) for both birds and mammals. Since the animal must be destroyed in the experiment, an efficient design is called for. For the birds, previous data indicate that $\log E[Y]$ and $X=\log Z$ are linearly related over $x\in[a,b]=[-5,3]$, with intercept $\theta_0\approx 0$ and slope $\theta_1\approx 1$. We thus take $E_0=0$, $E_1=1$ and $E_2=0$. The preceding theory (Theorem 1(II)) applies and shows that the canonical moments p_1^*, p_2^* of the Bayesian D-optimal design ξ^* within Ξ_2 are the solutions to

$$\frac{1}{p_2 p_1}=8, \quad \frac{1}{q_1}=8.$$

Table 3. Comparative values of the efficiency function for various designs.

	Prior		
	$\pi_1(\theta)$	$\pi_2(\theta)$	$\pi_3(\theta)$
ξ^*	4	4	4
ξ_1	0.773	0.773	0.773
ξ_2	3.126	2.966	2.901
ξ_3	3.561	3.309	3.217

Thus $p_1^* = 7/8, p_2^* = 1/7$. By (16), we have that the Bayesian D-optimal design within Ξ_2 is $\xi^*(1) = 1/2 = \xi^*(3)$. The practical interpretation is that half of the animals are to be chosen with log skeletal mass as close as possible to $x = 1$, with the other half being as close as possible to $x = 3$.

How does such a design compare with more *ad hoc* designs? To answer this we have computed values of the efficiency function $\Psi(\xi)$ for ξ^* and for three other designs: $\xi_1 = \frac{1}{2}(\delta_{-5} + \delta_3)$ (the classical D-optimal design), $\xi_2 = \frac{1}{3}(\delta_{-5} + \delta_{-1} + \delta_3)$ (the uniform design with three equally spaced support points) and $\xi_3 = \frac{1}{4}(\delta_{-5} + \delta_{-7/3} + \delta_{1/3} + \delta_3)$ (the uniform design with four equally spaced support points). The prior distributions considered all have $E_0 = 0$, $E_1 = 1$ and $E_2 = 0$, and are $\pi_1(\theta) = \frac{1}{2}(\delta_{(0,0.2,0)} + \delta_{(0,1.8,0)})$, $\pi_2(\theta) = \frac{1}{3}(\delta_{(0,0.2,0)} + \delta_{(0,1,0)} + \delta_{(0,1.8,0)})$, and $\pi_3(\theta) = \frac{1}{5}(\delta_{(0,0.2,0)} + \delta_{(0,0.5,0)} + \delta_{(0,1,0)} + \delta_{(0,1.5,0)} + \delta_{(0,1.8,0)})$. The efficiencies are exhibited in Table 3, and reveal the substantial savings in resources to be realized by the optimal design.

Table 4. Some examples of the Bayesian D-optimal designs within Ξ_3 when E_2 is small and positive.

E_2	$(p_1^*, p_2^*, p_3^*, p_4^*)$	(x_1^*, x_2^*, x_3^*)[a]	$\Psi(\xi)$
0.01	(0.744, 0.351, 0.803, 0.153)	(−1.879, 1.740, 3)	2.829
0.04	(0.699, 0.486, 0.715, 0.163)	(−2.974, 1.708, 3)	3.958
0.045	(0.616, 0.772, 0.236, 0.499)	(−4.839, 1.626, 3)	3.976
0.046	(0.609, 0.800, 0.118, 1)	(−5, 1.622, 3)	3.978
0.05	(0.610, 0.802, 0.116, 1)	(−5, 1.634, 3)	3.978
0.1	(0.616, 0.819, 0.100, 1)	(−5, 1.790, 3)	3.969

[a]Designs place mass 1/3 at each point x_i^*.

The experimenter may well wish for the protection of a three-point design such as ξ_2. Such a design, which allows for the fitting of a quadratic response or the testing of the fit of a linear response, is obtained if we take $k = n = 2$ in Theorem 1. With $E_0 = 0$, $E_1 = 1$ as above, but various non-zero values of E_2, we obtain the designs in Table 4. The efficiencies under the prior $\pi_2(\theta)$ are also given. We note that for $E_2 \geq 0.04$, the three point designs are very nearly as efficient as ξ^*.

ACKNOWLEDGMENTS

The research of Douglas Wiens is supported by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Chaloner, K. (1993). A note on optimal Bayesian design for nonlinear problems. *J. Stat. Planning and Inference* 37:229–235.

Chaloner, K., Clyde, M. (1996). The equivalence of constrained and weighted designs in multiple objective design problems. *J. Am. Stat. Assoc.* 91:1236–1244.

Chaloner, K., Verdinelli, I. (1995). Bayesian experimental design: A review. *Stat. Sci.* 10:273–304.

Dette, H., Neugebauer, H. M. (1996). Bayesian optimal one point designs for one parameter nonlinear models. *J. Stat. Planning and Inference* 52:17–31.

Dette, H., Neugebauer, H. M. (1997). Bayesian *D*-optimal designs for exponential models. *J. Stat. Planning and Inference* 60:331–350.

Dette, H., Studden, W. J. (1997). *The Theory of Canonical Moments with Applications in Statistics, Probability, and Analysis*. New York: Wiley.

Dette, H., Wong, W. K. (1996). Optimal Bayesian designs for models with partially specified heteroscedastic structure. *Ann. Stat.* 24:2108–2127.

Dette, H., Wong, W. K. (1998). Bayesian *D*-optimal designs on a fixed number of design points for heteroscedastic polynomial models. *Biometrika* 85:869–882.

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

Griffiths, D., Sandland, R. (1984). Fitting generalized allometric models to multivariate growth data. *Biometrics* 40:139–150.

Imhof, L. A. (2001). Maximin designs for exponential growth models and heteroscedastic polynomial models. *Ann. Stat.* 29:561–576.
Karlin, S., Studden, W. J. (1966). Optimal experimental designs. *Ann. Math. Stat.* 37:783–815.
MacDuffee, C. C. (1962). *Theory of Equations*. New York: Wiley.
Mukhopadhyay, S., Haines, L. M. (1995). Bayesian *D*-optimal designs for exponential growth model. *J. Stat. Planning and Inference* 44:385–397.
Prange, H. D., Anderson, J. F., Rahn, H. (1979). Scaling of skeletal to body mass in birds and mammals. *Am. Nat.* 113:103–122.
Pukelsheim, F. (1993). *Optimal Design of Experiments*. Toronto: John Wiley.
Sinha, S., Wiens, D. P. (2002). Robust sequential designs for nonlinear regression. *Canadian J. Stat.* 30:601–618.
Sokal, R. R., Rohlf, F. J. (1995). *Biometry: The Principles and Practice of Statistics in Biological Research*. New York: Freeman.
Szegö, G. (1959). *Orthogonal Polynomials*. Vol. XXIII, New York: American Mathematical Society Colloquium Publications.