The Annals of Statistics
2014, Vol. 42, No. 4, 1635–1656
DOI: 10.1214/14-AOS1241

E-OPTIMAL DESIGNS FOR SECOND-ORDER RESPONSE SURFACE MODELS

BY HOLGER DETTE[1,2] AND YURI GRIGORIEV[1]

Ruhr-Universität Bochum and St.-Petersburg State Electrotechnical University

E-optimal experimental designs for a second-order response surface model with $k \geq 1$ predictors are investigated. If the design space is the k-dimensional unit cube, Galil and Kiefer [*J. Statist. Plann. Inference* **1** (1977a) 121–132] determined optimal designs in a restricted class of designs (defined by the multiplicity of the minimal eigenvalue) and stated their universal optimality as a conjecture. In this paper, we prove this claim and show that these designs are in fact E-optimal in the class of all approximate designs. Moreover, if the design space is the unit ball, E-optimal designs have not been found so far and we also provide a complete solution to this optimal design problem.

The main difficulty in the construction of E-optimal designs for the second-order response surface model consists in the fact that for the multiplicity of the minimum eigenvalue of the "optimal information matrix" is larger than one (in contrast to the case $k = 1$) and as a consequence the corresponding optimality criterion is not differentiable at the optimal solution. These difficulties are solved by considering nonlinear Chebyshev approximation problems, which arise from a corresponding equivalence theorem. The extremal polynomials which solve these Chebyshev problems are constructed explicitly leading to a complete solution of the corresponding E-optimal design problems.

1. Introduction. Response surface methodology has become a standard tool in the analysis of experimental data. These models are used to study the influence of several input factors on a response variable by approximating complex functional relationships by "simple" linear or quadratic multivariate polynomial regression models, which are usually denoted as first or second-order response surface models [see, e.g., Myers, Montgomery and Anderson-Cook (2009)]. Numerous authors have worked on the construction of efficient and optimal experimental designs for response surface models. For first-order models, 2^k factorial and fractional factorial 2^{k-p} designs of resolution III are optimal with respect to

Received March 2014; revised May 2014.

[1]Supported by the Deutsche Forschungsgemeinschaft (SFB 823: Statistik nichtlinearer dynamischer Prozesse, Teilprojekt C2).

[2]Partially supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM107639.

MSC2010 subject classifications. Primary 62K05; secondary 41A50.

Key words and phrases. Response surface models, optimal designs, E-optimality, extremal polynomial, duality, nonlinear Chebyshev approximation.

the D-, G- and I-optimality criteria [see Anderson-Cook, Borror and Montgomery (2009)]. On the other hand, for the second-order response surface model the situation is more complicated and intuitively reasonable designs with a "simple" structure such as central composite designs are not optimal.

For this model, approximate designs in the sense of Kiefer (1974) have been investigated by several authors, where the methodology and optimal designs differ by the design space and optimality criterion under consideration (typical a k-dimensional cube, ball or simplex). D-optimal approximate designs for the second-order polynomial regression model on the ball and cube have been determined explicitly by Kiefer (1959, 1961b), Kiefer and Wolfowitz (1959), Kôno (1962), Farrell, Kiefer and Walbran (1967) [see also Rafajlowicz and Myszka (1988), Lim and Studden (1988) and Dette and Röder (1997) who determined optimal product designs for multivariate polynomial regression models in more general situations]. In particular, it is shown that D-optimal designs on a ball are at the same time rotatable designs. Considerably less attention has been paid to other optimality criteria. Laptev (1974), Denisov and Popov (1976) and Golikova and Pantchenko (1977) investigated A- and Q-optimal designs numerically, Galil and Kiefer (1977b) determined numerically rotatable optimal designs for the second-order response surface model, while Draper, Heiligers and Pukelsheim (2000) and Draper and Pukelsheim (2003) investigated optimal design problems in second-order mixture models. On the other hand, the explicit determination of optimal designs in the class of all approximate designs with respect to other criteria than the D-criterion seems to be a very hard problem, which has only been solved in rare circumstances.

In this paper, we study E-optimal designs for the second-order response surface models on the k-dimensional cube and ball. Among Kiefer‘s Φ_p-criteria [see Kiefer (1974)] the E-optimality criterion is not differentiable if the multiplicity of the minimum eigenvalue of the information matrix of the optimal design is larger than 1. This property makes the determination of E-optimal designs to an extremely hard and challenging problem. In fact, an analytical construction of E-optimal designs for linear regression models is very difficult and has only been achieved in the one-dimensional case for a limited number of linear and nonlinear models [see Melas (1982), Dette (1993), Pukelsheim and Studden (1993), Dette and Haines (1994), among others]. For models with more than one predictor, results can only be found sporadically in the literature. For example, Cheng (1987) and Dette and Studden (1993) identified E-optimal spring balance and chemical balance weighing designs. Galil and Kiefer (1977a) considered the second-order response surface model on the cube with k predictors and determined the E-optimal designs in the class of all designs, for which the corresponding information matrix has a minimum eigenvalue of multiplicity $k(k+1)/2$. However, to our best knowledge, the answer to the question, if these designs are in fact E-optimal in the class of all designs is still open. For the ball, the situation is even worse, and only E-optimal designs in the class of all rotatable designs are available [see, e.g., Galil and Kiefer (1977b)]. These designs are in fact not globally optimal and the determination of E-optimal designs for the second-order response surface model on the ball is an open and challenging problem.

The goal of the present paper is to provide complete answers to these questions and to characterize the structure and properties of E-optimal designs for the second-order response surface model. Our approach relies on a specific duality result for E-optimal designs, which relates the optimal design problem to a nonlinear Chebyshev approximation problem [see Melas (1982, 2006) or Pukelsheim (2006)]. In the dual problem, one has to determine a nonnegative polynomial with minimal sup-norm in a specific class of nonnegative (multivariate) polynomials, that is,

$$\mathcal{P} = \{ f^T(x) Z f(x) \,|\, \mathrm{trace}(Z) = 1; Z \geq 0 \}, \tag{1.1}$$

where x denotes the k-dimensional predictor, $f(x)$ is the vector of regression functions in the second-order response surface model and Z is a nonnegative definite matrix of appropriate dimension. This Chebyshev approximation problem is nonlinear and, therefore, extremely hard to solve explicitly. For the solution of the E-optimal design problem, this "optimal" polynomial, which is called extremal polynomial throughout this paper, will be constructed explicitly in Sections 3 and 4 if the design space is the cube and ball, respectively. As a consequence, we are able to provide a complete solution of these E-optimal design problems. In general, there exist several E-optimal designs which usually have a large number of support points. For this reason, particular attention is paid to the problem of constructing E-optimal designs with a small number of support points.

2. Optimal designs for response surface models. We consider the common linear regression model of the form

$$\mathbb{E}(Y|x) = f^T(x)\theta, \tag{2.1}$$

where Y denotes the (one-dimensional) response and the explanatory variable x varies in a compact design space, say $\mathcal{X} \subset \mathbb{R}^k$. In (2.1), the vector $f(x) = (f_1(x), \dots, f_m(x))^T \in \mathbb{R}^m$ is the vector of regression functions and $\theta = (\theta_1, \dots, \theta_m)^T \in \mathbb{R}^m$ denotes a vector of unknown parameters. We assume that N independent observations are available according to the model (2.1) where at each experimental condition x the response y is a realization of a normal distributed random variable Y with expectation given by (2.1) and (constant) variance $\sigma^2 > 0$. An approximate design in the sense of Kiefer (1974) is defined as probability measure on the design space $\mathcal{X}$ with finite support. The support points, say $x_{(1)}, \dots, x_{(n)}$, of an approximate design ξ define the locations where observations are taken, while the weights give the corresponding relative proportions of total observations to be taken at these points. If the design ξ has masses $\omega_i > 0$ at the different points $x_{(i)}$ $(i = 1, \dots, n)$ and N observations can be made by the experimenter, the quantities $\omega_i N$ are rounded to integers, say N_i, satisfying $\sum_{i=1}^{n} N_i = N$, and the experimenter takes N_i observations at each location $x_{(i)}$ $(i = 1, \dots, n)$. The information matrix of an approximate design ξ is defined by

$$M(\xi) = \int_{\mathcal{X}} f(x) f^T(x)\, d\xi(x) \in \mathbb{R}^{m\times m}, \tag{2.2}$$

and it is well known [see Jennrich (1969)] that under appropriate assumptions of regularity [in particular $\det(M(\xi)) > 0$ and $\lim_{N_i,N\to\infty} N_i/N = \omega_i > 0$; $i = 1, \dots, n$] the covariance matrix of the least squares estimator is approximately given by $\sigma^2 M^{-1}(\xi)/N$, where N denotes the total sample size.

Optimal designs maximize an appropriate statistical meaningful functional, say Φ, of the information matrix. Among the numerous criteria which have been proposed in the literature for this purpose [see Silvey (1980), Pázman (1986) or Pukelsheim (2006) among others], we consider in this paper the E-optimality criterion

$$\Phi_{-\infty}(\xi) = \lambda_{\min}\big(M(\xi)\big). \tag{2.3}$$

This criterion arises as a special case of Kiefer's Φ_p-optimality criteria, which are defined for $p \in (-\infty, 1]$ as

$$\Phi_p(M) = \big[m^{-1}\operatorname{tr}\big(M^p(\xi)\big)\big]^{1/p} = \left(m^{-1}\sum_{i=1}^{m}\lambda_i^p\big(M(\xi)\big)\right)^{1/p}, \tag{2.4}$$

that is $\Phi_{-\infty}(\xi) = \lim_{p\to-\infty}\Phi_p(\varepsilon)$ [see Kiefer (1974)]. In equation (2.4), the quantities $\lambda_1(M(\xi)), \dots, \lambda_m(M(\xi))$ denote the eigenvalues of the information matrix $M(\xi)$ and $\lambda_{\min}(M(\xi))$ its corresponding minimum eigenvalue. In contrast to the Φ_p-criteria with $p \in (-\infty, 1]$ the E-optimality criterion is not differentiable if the multiplicity of the minimum eigenvalue of the matrix $M(\xi)$ is larger than 1 and this property makes the determination of E-optimal designs to an extremely hard problem. In fact, E-optimal designs have been determined for a limited number of linear and nonlinear regression models [see the references cited in the Introduction]. An important tool for the determination of E-optimal designs is the following equivalence theorem which has been proved by several authors [see Melas (1982) or Pukelsheim (2006), e.g.].

THEOREM 2.1. *Let ξ^* denote a design and $\lambda_{\min}(M(\xi^*))$ the minimum eigenvalue of the information matrix $M(\xi^*)$ with multiplicity s. The design ξ^* is E-optimal if and only if there exist orthonormal eigenvectors $q_0, \dots, q_{s-1}$ of the matrix $M(\xi^*)$ corresponding to $\lambda_{\min}(M(\xi^*))$ and nonnegative weights $w_0, \dots, w_{s-1}$ with sum 1 such that the "extremal polynomial"*

$$\begin{aligned} d(x,\xi) &= f^T(x)(q_0, \dots, q_{s-1})\operatorname{diag}(w_0, \dots, w_{s-1})(q_0, \dots, q_{s-1})^T f(x) \\ &= \sum_{i=0}^{s-1} w_i\big(f^T(x)q_i\big)^2 \end{aligned}$$

satisfies for all $x \in \mathcal{X}$ *the inequality*

$$d(x,\xi) \le \lambda_{\min}\big(M(\xi^*)\big). \tag{2.5}$$

Moreover, the maximum on the left-hand side of (2.5) *is attained at the support points of the* E*-optimal design* ξ^*.

REMARK 2.1. It follows from general equivalence theory developed in convex design theory [see Pukelsheim (2006)] that there exists a duality between the E-optimal design problem and a nonlinear Chebyshev approximation problem, that is,

$$\max_{\xi} \lambda_{\min}\big(M(\xi)\big) = \min_{P_Z \in \mathcal{P}} \max_{x \in \mathcal{X}} \big|P_Z(x)\big|, \tag{2.6}$$

where $\mathcal{P} = \{P_Z(x) = f^T(x) Z f(x) | Z \in \mathbb{R}^{m\times m}, Z \ge 0, \operatorname{trace}(Z) = 1\}$ denotes a subset of the nonnegative "polynomials." In fact, if there is equality in (2.6) for a pair (ξ^*, Z^*), then ξ^* is an E-optimal design and P_{Z^*} a solution of the nonlinear Chebyshev approximation problem. This explains the name "extremal polynomial" in Theorem 2.1.

The second-order response surface model with a k-dimensional predictor appears as a special case of model (2.1), that is,

$$\mathbb{E}[Y|x] = \sum_{\|\alpha\|_1=0}^{2} \theta_\alpha x^\alpha, \tag{2.7}$$

where $\alpha = (\alpha_1, \dots, \alpha_k)^T \in \{0,1,2\}^k$ is a multiindex $x^\alpha = x^{\alpha_1} \cdots x^{\alpha_k}$ and $\|\alpha\|_1 = \alpha_1 + \cdots + \alpha_k$. In this case, the corresponding vector of regression function in the general linear model (2.1) is given by

$$f(x) = \big(1, x_1^2, \dots, x_k^2, x_1, \dots, x_k, x_1x_2, \dots, x_{k-1}x_k\big) \in \mathbb{R}^m, \tag{2.8}$$

where $m = \frac{(k+1)(k+2)}{2}$, $x = (x_1, \dots, x_k)^T$. In the following section, we consider optimal designs for the second-order regression model (2.7), where the design spaces are the unit ball with respect to the maximum norm $\|\cdot\|_\infty$ and the Euclidean norm $\|\cdot\|_2$, that is,

$$\begin{aligned} \mathcal{X} &= \mathcal{B}_\infty(1) := \big\{x \in \mathbb{R}^k \,|\, \|x\|_\infty \le 1\big\}, \\ \mathcal{X} &= \mathcal{B}_2(1) := \big\{x \in \mathbb{R}^k \,|\, \|x\|_2 \le 1\big\}. \end{aligned} \tag{2.9}$$

It turns out that designs with certain symmetry properties play a particular role for the construction of E-optimal designs. Throughout this paper, we call a design *symmetric* if for any $(\alpha_1, \dots, \alpha_k) \in \{0,1,2\}^k$ with $\|\alpha\|_1 = |\alpha_1| + \cdots + |\alpha_k| \le 2$ the moments

$$\int_{\mathcal{X}} x_1^{\alpha_1}, \dots, x_k^{\alpha_k} \xi(dx)$$

are invariant with respect to all permutations of $\alpha_1, \ldots, \alpha_k$ and vanish if there is at least one odd index among $\alpha_1, \ldots, \alpha_k$. In the following discussion, let $I_\ell \in \mathbb{R}^{\ell \times \ell}$ denote the identity matrix and $1_\ell = (1, \ldots, 1)^T \in \mathbb{R}^\ell$ denotes the vector with all elements equal to 1, then a straightforward calculation shows that the information matrix of a symmetric design in model (2.7) is of the form

$$M(\xi) = \int_{\mathcal{X}} f(x) f^T(x) \xi(dx)$$
$$= \begin{pmatrix} 1 & a1_k^T & 0 & 0 \\ a1_k & H & 0 & 0 \\ 0 & 0 & aI_k & 0 \\ 0 & 0 & 0 & bI_{(k(k-1))/2} \end{pmatrix} \in \mathbb{R}^{m \times m}, \tag{2.10}$$

where $m = \frac{(k+1)(k+2)}{2}$, $H = H(c; b) = (c - b)I_k + b1_k 1_k^T \in \mathbb{R}^{k \times k}$ denotes a circulant matrix with diagonal and off-diagonal elements c and b, respectively, and the entries a, b and c in (2.10) are given by

$$a = \int_{\mathcal{X}} x_1^2 \xi(dx), \qquad b = \int_{\mathcal{X}} x_1^2 x_2^2 \xi(dx), \qquad c = \int_{\mathcal{X}} x_1^4 \xi(dx). \tag{2.11}$$

Designs with information matrix of the form (2.10) will serve as candidates for E-optimal designs. Consider, for example, the case $k = 1$, where model (2.7) reduces to the well-known one-dimensional quadratic regression model $\theta_0 + \theta_1 x^2 + \theta_2 x$. If the designs space is given by $\mathcal{X} = [-1, 1]$ and the design ξ puts masses 1/5, 1/5 and 3/5 at the points -1, 1 and 0, respectively, the corresponding information matrix is given by

$$M(\xi^*) = \begin{pmatrix} 1 & \frac{2}{5} & 0 \\ \frac{2}{5} & \frac{2}{5} & 0 \\ 0 & 0 & \frac{2}{5} \end{pmatrix}.$$

It was shown by Kiefer (1974) that this design is in fact E-optimal for the univariate quadratic regression model and the minimum eigenvalue $\lambda_{\min} = \frac{1}{5}$ has multiplicity $s = 1$. For a similar statement in the univariate polynomial regression model of degree $d \geq 2$, see Pukelsheim and Studden (1993).

However, in the case $k \geq 2$, the multiplicity of the minimum eigenvalue of the matrix (2.10) is larger than 1 and as consequence the corresponding optimality criterion is not differentiable at the matrix $M(\xi)$ given by (2.10). This makes the determination of E-optimal designs substantially more difficult. For example, Galil and Kiefer (1977a) determined the E-optimal design on the cube $\mathcal{B}_\infty(1)$ in the subclass of all designs with information matrix of the form (2.10), where its minimum eigenvalue has multiplicity $\frac{k(k+1)}{2}$ (these calculations will be briefly presented at the beginning of the following section). To our best knowledge, the question, if the solution obtained by these authors in the restricted class yields in fact an E-optimal design for the second-order response surface model in the class of all approximate designs on the cube, has not been answered. Moreover, the E-optimal design problem for second-order regression models seems to be completely unsolved if the design space is given by the unit ball $\mathcal{B}_2(1)$.

In the following two sections, we will present a complete solution to these problems. For this purpose, we proceed in the following sections in two steps:

(I) In a first step, a candidate for the E-optimal design in the class of all designs with information matrix of the form (2.10) is identified. If the design space is given by the cube, our arguments coincide with those of Galil and Kiefer (1977a) and are presented here for the sake of completeness.

(II) In a second step, the E-optimality of the candidate design found by Galil and Kiefer (1977a) is proved by an application of Theorem 2.1. This requires the determination of an appropriate basis of the eigenspace corresponding to the minimum eigenvalue of $M(\xi)$ and the construction of the corresponding extremal polynomial in (2.6).

The E-optimal designs for the second-order response surface model will be identified in terms of the masses that they assign to specific sets which depend on the design space under consideration. Because in many applications it is desirable to obtain optimal designs with a minimal number of support points, we add a third step if the design space is the cube, that is,

(III) Identification of designs with a minimal number of support points.

3. E-optimal designs on the cube. In this section, we consider the second-order response surface model (2.1) on the design space $\mathcal{X} = \mathcal{B}_\infty(1) = [-1, 1]^k$. We start with a determination of a "good" candidate for an E-optimal symmetric design. Our arguments are similar to those given in Galil and Kiefer (1977a) and presented here for the sake of completeness (note that these authors only identified the candidate design and in the following we will prove its optimality in the class of all approximate designs). Observing the representation of the corresponding information matrix (2.10) the eigenvalues of the matrix $M(\xi)$ are given by a, b, and the eigenvalues by its upper $(k+1)\times(k+1)$ block,

$$M_{11}(\xi) = \begin{pmatrix} 1_k & a1_k^T \\ a1_k & H \end{pmatrix}, \tag{3.1}$$

where $H = H(c; b) = (c-b)I_k + b1_k1_k^T$. Define $D = [1 - c - (k-1)b]^2 + 4ka^2 > 0$, then all eigenvalues of the information matrix of a symmetric E-optimal design are given by

$$\begin{aligned} &\lambda_0 = \frac{1 + c + (k-1)b + \sqrt{D}}{2}, \qquad \lambda_1 = \frac{1 + c + (k-1)b - \sqrt{D}}{2}, \\ &\lambda_2 = \cdots = \lambda_k = c - b, \qquad \lambda_{k+1} = \cdots = \lambda_{2k} = a, \\ &\lambda_{2k+1} = \cdots = \lambda_m = b. \end{aligned} \tag{3.2}$$

Note that $\lambda_0 > \lambda_1$ and that λ_1 and λ_2 are increasing functions of c. Observing the identity

$$\det M(\xi) = a^k b^{k(k-1)/2}(c-b)^{k-1}\big[c+(k-1)b-ka^2\big] > 0$$

it is easy to see that the entries of a nonsingular matrix of the form (2.10) satisfy the inequalities

$$1 > \geq a \geq c > b > 0, \qquad c + b(k-1) > ka^2. \tag{3.3}$$

Therefore, we obtain $c = a$ and the problem of maximizing the minimum eigenvalue of $M(\xi)$ reduces to the maximization of

$$\begin{aligned}\lambda_{\min}\big(M(\xi)\big) &= \min\left\{\frac{1+a+(k-1)b-\sqrt{D}}{2}, a-b, a, b\right\}\\ &= \min\left\{\frac{1+a+(k-1)b-\sqrt{D}}{2}, a-b, b\right\},\end{aligned} \tag{3.4}$$

where the constant D is now represented as $D = [1-a-(k-1)b]^2 + 4ka^2$ and the second equality in (3.4) follows from $0 < a-b < a$ [see (3.3)]. We will now construct a candidate for the E-optimal design. Motivated by the solution of similar maximin problems, we suppose for this purpose that

$$\lambda_1 = \frac{1+a+(k-1)b-\sqrt{D}}{2} = a-b = b,$$

which gives $a = \frac{2}{5}$, $b = \frac{1}{5}$ as a unique (nontrivial) solution. This yields for the eigenvalues of the matrix $M(\xi)$

$$\begin{aligned}&\lambda_0 = 1 + \frac{k}{5}, \qquad \lambda_1 = \cdots = \lambda_k = \frac{1}{5},\\ &\lambda_{2k+1} = \cdots = \lambda_{(k(k-1))/2} = \frac{1}{5}, \qquad \lambda_{k+1} = \cdots = \lambda_{2k} = \frac{2}{5},\end{aligned} \tag{3.5}$$

where the corresponding multiplicities of $\lambda_0, \lambda_1, \lambda_{k+1}$ are given by 1, $\frac{k(k+1)}{2}$ and k, respectively. Hence, we obtain as a candidate for an E-optimal information matrix the matrix $M(\xi^*)$ in (2.10) with $a = c = \frac{2}{5}$, $b = \frac{1}{5}$, where the minimum eigenvalue is given by $\lambda_{\min}(M(\xi^*)) = \frac{1}{5}$. This means that the information matrix under consideration has a minimal eigenvalue with multiplicity $\frac{k(k+1)}{2} \geq 3$ whenever $k \geq 2$. The following result gives an answer to the question if the determined values for a and b yield in fact to an E-optimal information matrix.

THEOREM 3.1. *Any design ξ^* with an information matrix $M(\xi^*)$ of the form* (2.10) *and $a = c = \frac{2}{5}$ $b = \frac{1}{5}$ is E-optimal for the second-order response surface model* (2.7) *on the k-dimensional unit cube. In particular, Theorem* 2.1 *holds* *with*

$$d(x, \varepsilon) = \frac{1}{5}\left(1 - \frac{4}{k}\sum_{i=1}^{k} x_i^2(1 - x_i^2)\right). \tag{3.6}$$

The proof of Theorem 3.1 is complicated and deferred to Appendix A.1. Note that in contrast to the D-optimality criterion the optimal values for a and b do not depend on a dimension of the design space. This fact has been independently observed by Denisov and Popov (1976) and Galil and Kiefer (1977a), who identified the correct E-optimal information matrix but did not prove its optimality.

In the next step, we determine designs with corresponding information matrix specified in Theorem 3.1. For this purpose, we call a point $x \in \mathbb{R}^k$ a barycenter of depth $0 \le j \le k$ if j coordinates are equal to 0 and the remaining $k - j$ coordinates are equal to ± 1 [see Galil and Kiefer (1977a)]. The set of all barycenters of depth r is denoted E_r and for its cardinality we introduce the symbol

$$n_r := |E_r| = \binom{k}{r} 2^{k-r}, \qquad r = 0, 1, \ldots, k. \tag{3.7}$$

It was shown by Kiefer (1961a) and Farrell, Kiefer and Walbran (1967) that the support of every Φ_p-optimal design for the second-order response surface model on the cube is a subset of the set

$$E = \bigcup_{j=0}^{k} E_j. \tag{3.8}$$

Moreover, there always exists a symmetric optimal design. Throughout this section, we will describe these symmetric designs on the cube in terms of the $(k+1)$-dimensional vector $\xi = (\xi_0, \ldots, \xi_k)^T$, where ξ_i represents the mass assigned by the design to the set E_i of barycenters of depth i, that is $\xi_i = \xi(E_i)$ $(i = 0, \ldots, k)$. It turns out that there always exists an E-optimal design supported at most three sets E_i. For this purpose, we define for integers $0 \le r_1 < r_2 < r_3 \le k$ the matrix

$$A_{r_1,r_2,r_3} = \begin{pmatrix} 1 & 1 & 1 \\ \dfrac{k-r_1}{k} & \dfrac{k-r_2}{k} & \dfrac{k-r_3}{k} \\ \dfrac{k-r_1}{k}\dfrac{k-r_1-1}{k-1} & \dfrac{k-r_2}{k}\dfrac{k-r_2-1}{k-1} & \dfrac{k-r_3}{k}\dfrac{k-r_3-1}{k-1} \end{pmatrix}.$$

LEMMA 3.1. *There exists integers $0 \le r_1 < r_2 < r_3 \le k$ such that the system of linear equations*

$$A_{r_1,r_2,r_3}\xi = \left(1, \tfrac{2}{5}, \tfrac{1}{5}\right)^T \tag{3.9}$$

has a unique solution $\xi^* = (\xi_1^*, \xi_2^*, \xi_3^*)^T$ *satisfying* $\xi_i^* \geq 0$, $\sum_{i=1}^3 \xi_i^* = 1$. *Any design with masses*

$$\xi(E_{r_i}) = \xi_i^*, \qquad i = 1, 2, 3, \tag{3.10}$$

is E*-optimal for the second-order response surface model* (2.7).

PROOF. Let ξ denote a symmetric design and note that the moments in the matrix $M(\xi)$ defined in (2.10) have the representation

$$1 = \sum_{r=0}^{k} \xi_r, \qquad a = \sum_{r=0}^{k-1} a_r \xi_r, \qquad b = \sum_{r=0}^{k-2} b_r \xi_r, \tag{3.11}$$

where $\xi_r = \xi(E_r)$ is the measure of the set E_r of barycenters of depth r and

$$\begin{aligned} a_r &:= \binom{k-1}{r} 2^{k-r}, \qquad r \in \{0, \dots, k-1\}, \\ b_r &:= \binom{k-2}{r} 2^{k-r}, \qquad r \in \{0, \dots, k-2\}. \end{aligned} \tag{3.12}$$

By (3.11) and a remark on page 124 of Galil and Kiefer (1977a), there exist symmetric design ξ and three sets E_{r_1}, E_{r_2} and E_{r_3} such that (3.11) is satisfied for $a = \frac{2}{5}$ and $b = \frac{1}{5}$. A simple calculation shows that in this case the system of equations in (3.11) is equivalent to (3.9), which has a unique solution because $\det(A) = \frac{(r_1 - r_2)(r_1 - r_3)(r_2 - r_3)}{k^2(k-1)} \neq 0$. $\square$

It should be noted that not any solution of (3.9) will yield a vector of admissible weights $(\xi_{r_1}, \xi_{r_2}, \xi_{r_3}) = (\xi(E_{r_1}), \xi(E_{r_2}), \xi(E_{r_3}))$ (some components could be negative). Moreover, in general there exist many triples (r_1, r_2, r_3), such that the system (3.9) has a solution with nonnegative components and any such triple yields to at least one symmetric E-optimal design. For example, if (r_1, r_2, r_3) is such a triple with corresponding solution $(\xi(E_{r_1}), \xi(E_{r_2}), \xi(E_{r_3}))$ of (3.9), then a design ξ which assigns masses

$$\omega_{r_i, j} = \xi(\{x_{r_{i,j}}\}) = \frac{\xi(E_{r_i})}{n_{r_i}}; \qquad j = 1, \dots, n_{r_i}; i = 1, 2, 3;$$

to all points $x_{(r_i,1)} \cdots x_{(r_i, n_{r_i})} \in E_{r_i}$ is an E-optimal design for the second-order response surface model (2.7) on the unit cube $[-1, 1]^k$, where $n_j = \binom{k}{j} 2^{k-j}$ denotes the number of elements of the set E_j $(j = 0, \dots, k)$. The number of support points of such a design is given by

$$N(r_1, r_2, r_3) = \sum_{i=1}^{3} \binom{k}{r_i} 2^{k - r_i}$$

and usually rather large. For this reason, it is of interest to find designs with a minimal number of support points [see Farrell, Kiefer and Walbran (1967) or Pesotchinsky (1975)]. A reasonable approach to this problem is to look for E-optimal designs which are supported at only *two* sets of barycenters, say E_{r_1} and E_{r_2}. Because it can easily be shown that for a triple (r_1, r_2, r_3) with an admissible solution of (3.9) the weights $\xi(E_{r_i})$ are given by

$$\xi(E_{r_1}) = \frac{1}{5} \cdot \frac{2k^2 + k - 3k(r_2 + r_3) + 5r_2r_3}{(r_2 - r_1)(r_3 - r_1)}, \qquad i = 1, 2, 3, \tag{3.13}$$

it follows that symmetric E-optimal designs supported at only two sets of barycenters can be obtained from the *Diophantine equations*

$$2k^2 + k - 3k(s + t) + 5st = 0 \tag{3.14}$$

for $s, t = 0, \dots, k$. These equations have been solved numerically by Galil and Kiefer (1977a) if $k \le 25$ (see Table 1 in this reference). It should be pointed here that there does not always exist a solution of (3.14) (e.g., for $k = 2$, 6 or 8). Moreover, in general it is not clear that a solution of (3.14) necessarily yields to an E-optimal design with a minimal number of support points. For this reason, we display in Table 1 the E-optimal symmetric designs with a minimal number of support points for second-order response surface models with $k \le 24$ predictors. For example, if $k = 5$, the design with a minimal number of support points in

TABLE 1

Symmetric E-optimal designs with a minimal number of support points for second-order response surface models with $k \le 24$ predictors

k	(r_1, r_2, r_3)	$\xi(E_{r_1})$	$\xi(E_{r_2})$	$\xi(E_{r_3})$	k	(r_1, r_2, r_3)	$\xi(E_{r_1})$	$\xi(E_{r_2})$	$\xi(E_{r_3})$
1	(0, 1, –)	$\frac{2}{5}$	$\frac{3}{5}$	–	13	(0, 9, –)	$\frac{2}{15}$	$\frac{13}{15}$	–
2	(0, 1, 2)	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$	14	(0, 9, 14)	$\frac{25}{225}$	$\frac{182}{225}$	$\frac{18}{225}$
3	(–, 1, 3)	–	$\frac{3}{5}$	$\frac{2}{5}$	15	(0, 10, 15)	$\frac{3}{25}$	$\frac{21}{25}$	$\frac{1}{25}$
4	(0, 3, –)	$\frac{1}{5}$	$\frac{4}{5}$	–	16	(0, 11, –)	$\frac{7}{55}$	$\frac{48}{55}$	–
5	(0, 3, 5)	$\frac{2}{15}$	$\frac{10}{15}$	$\frac{3}{15}$	17	(0, 11, 17)	$\frac{18}{165}$	$\frac{136}{165}$	$\frac{11}{165}$
6	(0, 4, 6)	$\frac{3}{20}$	$\frac{15}{20}$	$\frac{2}{20}$	18	(0, 12, 18)	$\frac{7}{60}$	$\frac{51}{60}$	$\frac{2}{60}$
7	(0, 5, –)	$\frac{4}{25}$	$\frac{21}{25}$	–	19	(0, 13, –)	$\frac{8}{65}$	$\frac{57}{65}$	–
8	(0, 5, 8)	$\frac{9}{75}$	$\frac{56}{75}$	$\frac{10}{75}$	20	(0, 13, 20)	$\frac{49}{455}$	$\frac{380}{455}$	$\frac{26}{455}$
9	(0, 6, 9)	$\frac{2}{15}$	$\frac{12}{15}$	$\frac{1}{15}$	21	(0, 14, 21)	$\frac{4}{35}$	$\frac{30}{35}$	$\frac{1}{35}$
10	(0, 7, –)	$\frac{1}{7}$	$\frac{6}{7}$	–	22	(0, 15, –)	$\frac{3}{25}$	$\frac{22}{25}$	–
11	(0, 7, 11)	$\frac{8}{70}$	$\frac{55}{70}$	$\frac{7}{70}$	23	(0, 15, 23)	$\frac{32}{300}$	$\frac{253}{300}$	$\frac{15}{300}$
12	(0, 8, 12)	$\frac{5}{40}$	$\frac{33}{40}$	$\frac{2}{40}$	24	(0, 16, 24)	$\frac{9}{80}$	$\frac{69}{80}$	$\frac{2}{80}$

TABLE 2

Conjecture for the structure of E-optimal designs with a minimal number of support points for second-order response surface models with $k = 1, 2$ and $k \geq 4$ predictors, where $k = 3q + l$ and $s = 2q + l$ and $l = 0, \pm 1$

$l = +1$			$l = 0$			$l = -1$		
$\xi(E_0)$	$\xi(E_s)$	$\xi(E_k)$	$\xi(E_0)$	$\xi(E_s)$	$\xi(E_k)$	$\xi(E_0)$	$\xi(E_s)$	$\xi(E_k)$
$\frac{1}{5} \cdot \frac{q+2}{2q+1}$	$\frac{3}{5} \cdot \frac{3q+1}{2q+1}$	0	$\frac{1}{5} \cdot \frac{q+1}{2q}$	$\frac{3}{5} \cdot \frac{3q-1}{2q}$	$\frac{1}{5q}$	$\frac{1}{5} \cdot \frac{q}{2q-1}$	$\frac{1}{5} \cdot \frac{(3q-1)(3q-2)}{q(2q-1)}$	$\frac{2}{5q}$

only two sets has $N(2, 5) = 81$ support points in the set E_2 and E_5 [see Galil and Kiefer (1977a)], while the design with the minimal number of $N(0, 3, 5) = 73$ support points in the sets E_0, E_3 and E_5.

REMARK 3.1. Based on our numerical results, we found a remarkable structure for the E-optimal designs with a minimal number of support points for the second-order response surface model with k predictors, whenever $k \neq 3$. The E-optimal design for the second-order response surface model with a minimal number of support points is always supported at the sets E_0 and E_k and a third set E_s. If $k = 3q + l$ where $l = 0, \pm 1$, then $s = 2q + l$. The particular structure is displayed in Table 2, which also contains the weights assigned by the E-optimal design to these sets.

EXAMPLE 3.1. Galil and Kiefer (1977a) presented in Table 2 of their paper E-optimal designs (obtained as limits of Φ_p-optimal designs as $p \to -\infty$). Note that not all designs in this class have the minimal number of support points. For example, if $k = 6$ the E-optimal design obtained by Galil and Kiefer (1977a) puts masses 0.040, 0.400 and 0.560 at the sets E_0, E_2 and E_5, respectively, and has 316 support points. The E-optimal design obtained from Table 2 puts masses $\xi(E_0) = 0.15$, $\xi(E_4) = 0.75$, $\xi(E_6) = 0.10$ and has only 125 support points.

4. E-optimal designs on the unit ball. In this section, we consider the E-optimal design problem for the second-order response surface model on the k-dimensional ball $\mathcal{B}_2(1) = \{x \in \mathbb{R}^k : \|x\|_2 \leq 1\}$. The general strategy for the solution of the optimal design problem will be similar as the one given for the cube and we start identifying a good candidate for the E-optimal design. If the design space is the ball, then the sets E_{r_i} of barycenters of depth r_i will be replaced by three sets F_0, F_{k-1} and F_k as candidate sets for the support of E-optimal designs. Here, F_0 consists of the 2^k vertices $x = (\pm\frac{1}{\sqrt{k}}, \dots, \pm\frac{1}{\sqrt{k}})^T \in \mathbb{R}^k$ of the cube $\mathcal{B}_\infty(1/\sqrt{k})$ inscribed in k-dimensional ball $\mathcal{B}_2(1)$, F_{k-1} consists of the centers $\pm e_i$ of the $(k-1)$-dimensional faces of $\mathcal{B}_\infty(1)$ [here $e_i = (0, \dots, 0, 1, 0, \dots, 0)^T$ denotes the

ith unit vector] and F_k contains only the center of the ball. Note that the cardinality of these sets are given by

$$|F_0| = 2^k, \qquad |F_{k-1}| = 2k, \qquad |F_k| = 1. \tag{4.1}$$

As a consequence, there is no necessity to search for the minimally supported designs on the unit ball.

Consider a symmetric design ξ which is supported on the sets F_0, F_{k-1} and F_k introduced in the previous paragraph. Its information matrix $M(\xi)$ in the second-order response surface model (2.1) is of the form (2.10) with corresponding eigenvalues given by (3.2) where $D = [1 - (c - b) - kb]^2 + 4ka^2 > 0$. Moreover, from the definition of ξ we have for the entries defined in the matrix (2.10)

$$\begin{aligned} a &= k^{-1}\xi(F_0) + k^{-1}\xi(F_{k-1}), \\ b &= k^{-2}\xi(F_0), \\ c &= k^{-2}\xi(F_0) + k^{-1}\xi(F_{k-1}), \end{aligned} \tag{4.2}$$

and it now follows that

$$\xi(F_{k-1}) = k(a - kb) = k(c - b). \tag{4.3}$$

Substituting this identity into expression (3.2) for λ_1 yields

$$\lambda_1 = \frac{1 + a - \sqrt{(1-a)^2 + 4ka^2}}{2}. \tag{4.4}$$

Therefore, the problem of determining an E-optimal (symmetric) design in the class of measures supported at the sets F_0, F_{k-1} and F_k reduces to the maximization of [note that $a > b$ because otherwise by (3.3) and (4.3) we would obtain $\xi(F_{k-1}) = 0$, hence $a = b = c$, which is impossible]

$$\lambda_{\min}(M(\xi)) = \min\left\{\frac{1 + a - \sqrt{(1-a)^2 + 4ka^2}}{2}, c - b, b\right\}, \tag{4.5}$$

where $0 \le a, b, c \le 1$. In order to construct a good candidate, say ξ^*, for the E-optimal information matrix we assume that for the optimal design all elements in (4.5) are identical, which yields by a straightforward calculation [observing (4.3)] for the elements in the matrix (2.10)

$$a = \frac{k+1}{k^2 + 2k + 2}, \qquad b = \frac{1}{k^2 + 2k + 2}, \qquad c = \frac{2}{k^2 + 2k + 2}. \tag{4.6}$$

In this case,

$$\lambda_{\min}(M(\xi^*)) = \frac{1}{k^2 + 2k + 2} \tag{4.7}$$

is the minimal eigenvalue of the matrix $M(\xi^*)$ with multiplicity $s = \frac{k(k+1)}{2}$. Since this solution has been obtained under the constraint that the designs is supported at the sets F_0, F_{k-1} and F_k and that all elements in (4.5) are identical, it is not clear that the resulting information matrix is in fact E-optimal. In a second step, we establish this optimality. In order to explain the general principle, we begin with an example.

EXAMPLE 4.1. Consider the second-order response surface model with $k=2$ predictors. Thus, we have $m=6$ regression functions, the minimum eigenvalue is given by $\lambda_{\min}=\frac{1}{10}$, with multiplicity $s=3$. For a corresponding orthogonal basis in Theorem 2.1, we choose

$$q_0=(2,-3,-3,0,0,0)^T,$$
$$q_1=(0,-1,1,0,0,0)^T,$$
$$q_2=(0,0,0,0,0,1)^T,$$

which yields $\|q_0\|^2=22$, $\|q_1\|^2=2$, $\|q_2\|^2=1$ and for the extremal polynomial

$$d(x,\varepsilon)=\frac{w_0}{\|q_0\|^2}\left(2-3\sum_{j=1}^{2}x_j^2\right)^2+\frac{w_1}{\|q_1\|^2}\big(x_1^2-x_2^2\big)^2+\frac{w_2}{\|q_2\|^2}(x_1x_2)^2. \tag{4.8}$$

The vector of weights w is identified by the condition that there must be equality in (2.5) for the support points of the E-optimal design and the condition $w_0+w_1+w_2=1$. Using the points $x_{(0)}=(0,0)^T\in F_0$ and $x_{(1)}=(1,0)^T\in F_1$, we obtain for the vector $w=(\frac{11}{20},\frac{3}{20},\frac{6}{20})$ and

$$d(x,\varepsilon)=\frac{1}{10}\left(1-3\left(\sum_{i=1}^{2}x_i^2\right)\left(1-\sum_{i=1}^{2}x_i^2\right)\right)=\frac{1}{10}\big(1-3\|x\|_2^2\big(1-\|x\|_2^2\big)\big).$$

Obviously, we have for all x with $\|x\|_2\le 1$

$$d(x,\varepsilon)\le\frac{1}{10}=\lambda_{\min}\big(M(\xi)\big),$$

and by Theorem 2.1 any design with information matrix of the form (2.10) with $a=\frac{3}{10}$, $b=\frac{1}{10}$, $c=\frac{2}{10}$ is E-optimal for the second-order response surface model on the ball.

The following result provides a similar statement in the general case. Its proof is complicated and therefore deferred to Appendix A.2.

THEOREM 4.1. *Let ξ^* denote a symmetric design on the ball $\mathcal{B}_2(1)$, which puts masses*

$$\begin{aligned}\xi(F_0)&=\frac{k^2}{k^2+2k+2},\\ \xi(F_{k-1})&=\frac{k}{k^2+2k+2},\\ \xi(F_k)&=\frac{k+2}{k^2+2k+2}\end{aligned}\tag{4.9}$$

at the sets F_0, F_{k-1} *and* F_k, *respectively, then* ξ^* *is* E-*optimal for the second-order response surface model on the* k-*dimensional unit ball. Moreover, the minimal eigenvalue of the matrix* $M(\xi^*)$ *is given by* (4.7) *with multiplicity* $s = \frac{k(k+1)}{2}$ *and the extremal polynomial in Theorem* 2.1 *can be chosen as*

$$d(x,\varepsilon) = \frac{1}{k^2+2k+2}\left\{1 - \frac{2(k+1)}{k}\|x\|_2^2\bigl(1-\|x\|_2^2\bigr)\right\}. \tag{4.10}$$

We conclude this section with a brief discussion of rotatable designs, which are defined as designs for which the dispersion function $U : \mathcal{B}_2(1) \to \mathbb{R}; x \to U(x,\xi) = f^T(x)M^{-1}(\xi)f(x)$ is invariant with respect to orthogonal transformations, that is,

$$U(x,\xi) = U(Ox,\xi) \qquad \forall x \in \mathbb{R}^k, \tag{4.11}$$

whenever O is an orthogonal $k \times k$ matrix. Note that this property is equivalent to the fact that the function $U(x,\xi)$ depends only of the radius $\|x\|_2$. The following result characterizes the rotatability of a symmetric design with information matrix of the form (2.10) and will be used to investigate if E-optimal designs in the class of all rotatable designs are also E-optimal in the class of all symmetric designs.

LEMMA 4.1. *Let* ξ *denote a symmetric design on the ball* $\mathcal{B}_2(r)$ *of radius* $r > 0$ *with information matrix of the form* (2.10). *Then the design* ξ *is rotatable for the second-order response surface model, if and only if the condition*

$$c = 3b \tag{4.12}$$

is satisfied. Moreover, the uniform distribution on sphere $\partial\mathcal{B}_2(r)$ *denoted by* $\mathcal{U}(\partial\mathcal{B}_2(r))$ *defines a rotatable design.*

PROOF. Let ξ denote a design with information matrix (2.10). A simple calculation shows that the inverse of the $k \times k$ upper block (3.1) of the matrix $M(\xi)$ is given by

$$M_{11}^{-1}(\xi) = \begin{pmatrix} \varkappa & q1_k^T \\ q1_k & G \end{pmatrix},$$

where $\varkappa = (c + b(k-1))/Q_0$, $q = -a/Q_0$, $Q_0 = c - b + (b - a^2)k$, and $G = (d-e)I_k + e1_k1_k^T$ is a circulant matrix with diagonal elements d and off-diagonal elements e defined by

$$e = \frac{a^2 - b}{(c-b)Q_0}, \qquad d = Q_0^{-1} - e(k-1),$$

respectively. As a consequence, we obtain for the function U the representation

$$\begin{aligned}U(x,\xi) &= f^T(x)M^{-1}(\xi)f(x)\\ &= \varkappa + (a^{-1}+2q)\|x\|_2^2 + (b^{-1}+2e)\sum_{i<j}^{k}(x_ix_j)^2 + d\sum_{i=1}^{k}x_i^4\\ &= \varkappa + (a^{-1}+2q)\|x\|_2^2 + \left(\frac{1}{2b}+e\right)\|x\|_2^4 + \left(d-e-\frac{1}{2b}\right)\sum_{i=1}^{k}x_i^4.\end{aligned}$$

Now the design is rotatable if and only if the function $U(x,\xi)$ depends only on the radius $\|x\|_2$, that is,

$$0 = d - e - (2b)^{-1} = (3b-c)/2b(c-b),$$

which proves the first part of the assertion. The second part follows by a straightforward calculation of the moments of the uniform distribution on the sphere $\partial\mathcal{B}_2(r)$. □

Galil and Kiefer (1977b) have determined the E-optimal rotatable designs on the ball $\mathcal{B}_2(r)$ for the second-order response surface model (2.7), which are given by

$$\xi^*(\alpha) = (1-\alpha)\xi(\{0\}) + \alpha\mathcal{U}(\partial\mathcal{B}_2(r)),$$

where the parameter α is defined by

$$\alpha = \begin{cases}\dfrac{k(k+1)(k+2)}{(k+1)r^4+k(k+2)^2}, & r^2 \le k+2,\\[2ex] \dfrac{k(r^2-1)}{r^2(r^2+k-1)}, & r^2 \ge k+2.\end{cases} \tag{4.13}$$

If the design space is given by the unit ball $\mathcal{B}_2(1)$ this design is not E-optimal in the class of all designs. In fact, the symmetric E-optimal design ξ^* determined in Theorem 4.1 does not satisfy condition (4.12) and is therefore not rotatable. The minimum eigenvalue of the matrix $M(\xi^*)$ is given by (4.7), while the minimum eigenvalue of the E-optimal design in the class of all rotatable designs is given by

$$\lambda_{\min}(M(\xi(\alpha))) = \frac{k+1}{k^3+4k^2+5k+1} < \frac{1}{k^2+2k+2} = \lambda_{\min}(M(\xi^*)).$$

We finally note that there exists a difference between the E- and D-optimality criterion with respect to the property of rotatability. In contrast to the E-optimal design, the D-optimal design for the second-order response surface model on the ball $\mathcal{B}_2(1)$ is also rotatable [see Kiefer (1961a)].

APPENDIX: PROOFS OF THEOREMS 3.1 AND 4.1

A.1. Proof of Theorem 3.1. Throughout the proof, we assume $k \geq 2$, the case $k = 1$ has been treated in Pukelsheim and Studden (1993), for example. Recall the definition of the vector of regression functions (2.8) in model (2.1) and note that for the optimal design ξ^* under consideration we have $a = \frac{2}{5}$ and $b = \frac{1}{5}$ in the matrix (2.10) with minimum eigenvalue given by $\lambda_{\min}(M(\xi^*)) = \frac{1}{5}$ (see the discussion at the beginning of Section 3). Consequently, a possible candidate $q_0, \ldots, q_{s-1}$ for the basis of the eigenspace corresponding to $\lambda_{\min}(M(\xi^*))$ is given by

$$Q = (q_0, \ldots, q_{s-1})$$

$$= \begin{pmatrix} G_{k\times(k+1)} & 0_{k\times k} & 0_{k\times((k(k-1))/2)} \\ 0_{((k(k-1))/2)\times(k+1)} & 0_{((k(k-1))/2)\times k} & I_{(k(k-1))/2} \end{pmatrix}^T, \tag{A.1}$$

with an appropriate matrix $G_{k\times(k+1)} \in \mathbb{R}^{k\times k+1}$ (here and throughout this section $0_{r\times s}$ denotes the matrix with all entries given by 0). This means that the unit vectors $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ are eigenvectors of the matrix $M(\xi^*)$ for $i = 2k + 2, \ldots, m = \frac{(k+1)(k+2)}{2}$. It turns out that it is reasonable to use a vector of weights, which is of the form

$$w = (w_0, w_1, \ldots, w_{k-1}, 0, \ldots, 0)^T \in \mathbb{R}^s \tag{A.2}$$

in Theorem 2.1. Observing (A.1), it then follows that for vectors of this type only the $k + 1$ functions

$$\{1, x_1^2, \ldots, x_k^2\}$$

will appear in the corresponding extremal polynomial. We now construct the remaining part of the orthogonal basis in (A.1) by choosing the block matrix

$$G_{k\times(k+1)} = \begin{pmatrix} k & -2 & -21_{k-1}^T \\ 0_{k-1} & -1_{k-1} & L \end{pmatrix} \in \mathbb{R}^{k\times k+1}, \tag{A.3}$$

where the matrix $L = (L_{ij})_{i,j=1}^{k-1} \in \mathbb{R}^{(k-1)\times(k-1)}$ is defined by

$$L_{ij} = \begin{cases} -1, & i + j < k, \\ k - i, & i + j = k, \\ 0, & i + j > k. \end{cases}$$

This gives for the eigenvectors $q_0, \ldots, q_{k-1}$ [defined by the first k rows of the matrix Q in (A.1)]

$$\|q_0\|^2 = k^2 + 4k, \qquad \|q_r\|^2 = (k - r)(k - r + 1), \qquad r = 1, \ldots, k - 1.$$

With the notation $b_i(x) = (q_i^T f(x))^2$, the extremal polynomial in Theorem 2.1 has the representation

$$d(x, \varepsilon) = \sum_{i=0}^{k-1} w_i b_i(x), \tag{A.4}$$

where we have used (A.2) and the function $b_0, \dots, b_{k-1}$ are given by

$$
\begin{aligned}
b_0(x) &= \left(k - 2\sum_{i=1}^{k} x_i^2\right)^2 \cdot \frac{1}{\|q_0\|^2}, \\
b_r(x) &= \left(\sum_{i=1}^{k-r} x_i^2 - (k-r)x_{k-r+1}^2\right)^2 \cdot \frac{1}{\|q_r\|^2}, \qquad r = 1, \dots, k-1.
\end{aligned} \tag{A.5}
$$

The coefficients w_i in the polynomial (A.4) are now determined by the condition $d(x,\xi) = \lambda_{\min}(M(\xi)) = \frac{1}{5}$ at the points $x^{(r)} = (0, \dots, 0, 1, \dots, 1)^T$ with $\|x^{(r)}\|_1 = r$ and the fact that $\sum_{i=0}^{k-1} w_i = 1$. This leads to the matrix equation

$$
B(w_0, \dots, w_{k-1})^T = J_0, \tag{A.6}
$$

where $J_0 = (\frac{1}{5}, \dots, \frac{1}{5}, 1)^T \in \mathbb{R}^k$ and the matrix $B = (B_{ir})_{i,r=0}^{k-1,k-1}$ is a lower triangular matrix with nonvanishing elements

$$
B_{ir} = \begin{cases}
\dfrac{k}{k+4}, & i = 0, r = 0, \\
\dfrac{(k-2i)^2}{k^2+4k}, & i = 1, \dots, k-2, r = 0, \\
\dfrac{(k-i)^2}{(k-r)(k-r+1)}, & i = 1, \dots, k-2, r = 1, \dots, i, \\
1, & i = k-1, r = 0, \dots, k-1.
\end{cases}
$$

A simple calculation shows $(w_0, \dots, w_{k-1})^T = B^{-1}J_0 = \frac{1}{5k}(k+4, 4, 4, \dots, 4) \in \mathbb{R}^k$ and $w = (\frac{k+4}{5k}, \frac{4}{5k}, \dots, \frac{4}{5k}, 0, \dots, 0)^T$ is the vector which will be used for the calculation of a candidate for the extremal polynomial. For this purpose, we introduce the notation

$$
\alpha_r = \frac{w_r}{\|q_r\|^2} = \begin{cases}
\dfrac{1}{5k^2}, & r = 0, \\
\dfrac{4}{5k(k-r+1)(k-r)}, & r = 1, \dots, k-1,
\end{cases} \tag{A.7}
$$

and a tedious but straightforward algebra yields for the polynomial (A.4) the representation

$$
\begin{aligned}
d(x,\varepsilon) &= \alpha_0\left(k - 2\sum_{i=1}^{k} x_i^2\right)^2 + \sum_{r=1}^{k-1} \alpha_r \left(\sum_{i=1}^{k-r} x_i^2 - (k-r)x_{k-r+1}^2\right)^2 \\
&= \frac{1}{5}\left(1 - \frac{4}{k}\sum_{i=1}^{k} x_i^2(1-x_i^2)\right),
\end{aligned}
$$

which coincides with (3.6). As a consequence, we obtain for all $x \in [-1, 1]^k$

$$d(x, \xi^*) \leq \lambda_{\min}(M(\xi^*)) = \tfrac{1}{5},$$

and by Theorem 2.1 the matrix $M(\xi^*)$ is an E-optimal information matrix.

A.2. Proof of Theorem 4.1. The proof proceeds in a similar way as the proof of Theorem 3.1 but differs in some essential details from it. To be precise, recall that for the design ξ^* under consideration the minimal eigenvalue of its information matrix $M(\xi^*)$ is given by $\lambda_{\min}(M(\xi^*)) = \frac{1}{k^2+2k+2}$ and has multiplicity $s = \frac{k(k+1)}{2}$. As in the proof of Theorem 3.1 we consider the matrix defined by (A.1) as a candidate for an orthonormal basis of the corresponding eigenspace. For the matrix $G_{k\times(k+1)} \in \mathbb{R}^{k\times k+1}$, we now use

$$G_{k\times(k+1)} = \begin{pmatrix} k & -(k+1) & -(k+1)1_{k-1}^T \\ 0_{k-1} & 1_{k-1} & L \end{pmatrix}, \tag{A.8}$$

where $L = (L_{ij})_{i,j=1}^{k-1} \in \mathbb{R}^{(k-1)\times(k-1)}$ is a lower triangular matrix with nonvanishing elements

$$L_{ij} = \begin{cases} -i, & i = j, \\ 1, & i > j. \end{cases} \tag{A.9}$$

Consequently, we have

$$\begin{aligned} \|q_0\|^2 &= k^2 + k(k+1)^2, \\ \|q_r\|^2 &= r(r+1), \qquad r = 1, \dots, k-1, \\ \|q_r\|^2 &= 1, \qquad r = k, \dots, s-1, \end{aligned}$$

and with the notation $b_i(x) := (f^T(x)q_i)^2$ the candidate for the extremal polynomial in (2.5) has the representation

$$d(x, \xi^*) = \sum_{i=0}^{s-1} w_i b_i(x), \tag{A.10}$$

where [recall the definition of the vector f in (2.8)]

$$\begin{aligned} b_0(x) &= \left(k - (k+1)\sum_{i=1}^{k} x_i^2\right)^2 \cdot \frac{1}{\|q_0\|^2}, \\ b_r(x) &= \left(\sum_{i=1}^{r} x_i^2 - r x_{r+1}^2\right)^2 \cdot \frac{1}{\|q_r\|^2}, \qquad r = 1, \dots, k-1, \\ b_{k-2+i+j}(x) &= (x_i x_j)^2, \qquad i = 1, \dots, k-1;\ j = i+1, \dots, k \end{aligned}$$

(note that the eigenvectors corresponding to $b_{k-2+i+j}$ satisfy $\|q_r\| = 1$).

For determination of coordinates of the vector w, we use again the fact that there must be equality in condition (2.5) of Theorem 2.1 for the support points of an E-optimal design. For the point $x_{(0)} = 0 \in \mathbb{R}^k$, the condition $d(x_{(0)}, \varepsilon) = \lambda_{\min}(M(\xi^*))$ and (A.10) then yields

$$w_0 = \frac{k^2 + 3k + 1}{k(k^2 + 2k + 2)}. \tag{A.11}$$

We now try to find a candidate for the remaining weights under the additional assumption that $p_1 := w_1 = \cdots = w_{k-1}$ and $p_2 := w_k = \cdots = w_{s-1}$. Because the sum of all weights is 1, this gives the equality

$$w_0 + (k-1)p_1 + \frac{k(k-1)}{2} p_2 = 1. \tag{A.12}$$

Finally, we use one more point $x_{(1)} = (1, 0, \ldots, 0)^T \in F_{k-1}$ in the condition $d(x_{(1)}, \xi^*) = \lambda_{\min}(M(\xi^*))$ to obtain the equation

$$w_0 + p_1 \sum_{r=1}^{k-1} \|q_r\|^{-2} = \lambda_{\min}(M(\xi^*)) = \frac{1}{k^2 + 2k + 2}. \tag{A.13}$$

Since $\sum_{r=1}^{k-1} \|q_r\|^{-2} = 1 - k^{-1}$, we finally obtain from (A.11)–(A.13) for the weights

$$\begin{aligned} w_0 &= \frac{k^2 + 3k + 1}{k(k^2 + 2k + 2)}, \\ w_1 = \cdots = w_{k-1} &= \frac{k+1}{k(k^2 + 2k + 2)}, \\ w_k = \cdots = w_{s-1} &= \frac{2(k+1)}{k(k^2 + 2k + 2)}. \end{aligned} \tag{A.14}$$

Substituting these expressions in (A.10) yields by a straightforward calculation

$$d(x, \xi^*) = \frac{1}{k^2 + 2k + 2}\left(1 - \frac{2(k+1)}{k} \|x\|_2^2 (1 - \|x\|_2^2)\right) \tag{A.15}$$

as a candidate for the extremal polynomial. Obviously, we have

$$d(x, \xi^*) \le \frac{1}{k^2 + 2k + 2} = \lambda_{\min}(M(\xi^*))$$

for all $x \in \mathcal{B}_2(1)$, and by Theorem 2.1 the information matrix $M(\xi^*)$ defined in (2.10) with moments (4.6) is E-optimal for the second-order response surface model on the ball.

Acknowledgements. This work was done during a visit of the second author (Y. Grigoriev) at the Department of Mathematics, Ruhr-Universität Bochum, Germany. The authors would like to thank M. Stein who typed this manuscript with considerable technical expertise and F. Pukelsheim for some useful hints regarding related literature. We are also very grateful to two unknown referees for their constructive comments on the first version of our paper. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

ANDERSON-COOK, C. M., BORROR, C. M. and MONTGOMERY, D. C. (2009). Response surface design evaluation and comparison. *J. Statist. Plann. Inference* **139** 629–641.

CHENG, C.-S. (1987). An application of the Kiefer–Wolfowitz equivalence theorem to a problem in Hadamard transform optics. *Ann. Statist.* **15** 1593–1603. MR0913576

DENISOV, V. I. and POPOV, A. A. (1976). $A - E$-optimal and orthogonal designs of experiments for polynomial models. Preprint. Scientific Council on a Complex Problem "Cybernetic," Academy of Science (in Russian). Moscow, Russia.

DETTE, H. (1993). A note on E-optimal designs for weighted polynomial regression. *Ann. Statist.* **21** 767–771. MR1232518

DETTE, H. and HAINES, L. M. (1994). E-optimal designs for linear and nonlinear models with two parameters. *Biometrika* **81** 739–754. MR1326423

DETTE, H. and RÖDER, I. (1997). Optimal discrimination designs for multifactor experiments. *Ann. Statist.* **25** 1161–1175. MR1447745

DETTE, H. and STUDDEN, W. J. (1993). Geometry of E-optimality. *Ann. Statist.* **21** 416–433. MR1212185

DRAPER, N. R., HEILIGERS, B. and PUKELSHEIM, F. (2000). Kiefer ordering of simplex designs for second-degree mixture models with four or more ingredients. *Ann. Statist.* **28** 578–590. MR1790010

DRAPER, N. R. and PUKELSHEIM, F. (2003). Canonical reduction of second-order fitted models subject to linear restrictions. *Statist. Probab. Lett.* **63** 401–410. MR1996188

FARRELL, R. H., KIEFER, J. and WALBRAN, A. (1967). Optimum multivariate designs. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (*Berkeley*, *Calif.*, 1965/66) 113–138. Univ. California Press, Berkeley, CA. MR0214248

GALIL, Z. and KIEFER, J. (1977a). Comparison of design for quadratic regression on cubes. *J. Statist. Plann. Inference* **1** 121–132. MR0518970

GALIL, Z. and KIEFER, J. (1977b). Comparison of rotatable designs for regression on balls. I. Quadratic. *J. Statist. Plann. Inference* **1** 27–40. MR0518969

GOLIKOVA, T. I. and PANTCHENKO, L. A. (1977). Continuous A and Q-optimal second order designs on a cube. In *Regression Experiments* (*Design and Analysis*) (*in Russian*) (V. V. Nalimov, ed.) 71–84. Moscow Univ., Moscow.

JENNRICH, R. I. (1969). Asymptotic properties of non-linear least squares estimators. *Ann. Math. Statist.* **40** 633–643. MR0238419

KIEFER, J. (1959). Optimum experimental designs. *J. Roy. Statist. Soc. Ser. B* **21** 272–319. MR0113263

KIEFER, J. (1961a). Optimum experimental designs. V. With applications to systematic and rotatable designs. In *Proc.* 4*th Berkeley Sympos. Math. Statist. and Prob.*, *Vol. I* 381–405. Univ. California Press, Berkeley, CA. MR0133941

KIEFER, J. (1961b). Optimum designs in regression problems. II. *Ann. Math. Statist.* **32** 298–325. MR0123408

KIEFER, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2** 849–879. MR0356386

KIEFER, J. and WOLFOWITZ, J. (1959). Optimum designs in regression problems. *Ann. Math. Statist.* **30** 271–294. MR0104324

KÔNO, K. (1962). Optimum design for quadratic regression on k-cube. *Mem. Fac. Sci., Kyushu Univ., Ser. A, Math.* **16** 114–122. MR0153090

LAPTEV, V. N. (1974). Some problems relating to construction of regression experimental designs by means of computer. Ph.D. thesis, Novosibirsk State Technical Univ., Russia.

LIM, Y. B. and STUDDEN, W. J. (1988). Efficient D_s-optimal designs for multivariate polynomial regression on the q-cube. *Ann. Statist.* **16** 1225–1240. MR0959198

MELAS, V. B. (1982). A duality theorem and E-optimality (translated from Russian). *Industrial Laboratory* **48** 295–296.

MELAS, V. B. (2006). *Functional Approach to Optimal Experimental Design*. Springer, New York. MR2193670

MYERS, R. H., MONTGOMERY, D. C. and ANDERSON-COOK, C. M. (2009). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 3rd ed. Wiley, Hoboken, NJ. MR2464113

PÁZMAN, A. (1986). *Foundations of Optimum Experimental Design*. Reidel, Dordrecht. MR0838958

PESOTCHINSKY, L. L. (1975). D-optimum and quasi-D-optimum second-order designs on a cube. *Biometrika* **62** 335–340. MR0398016

PUKELSHEIM, F. (2006). *Optimal Design of Experiments*. SIAM, Philadelphia, PA. MR2224698

PUKELSHEIM, F. and STUDDEN, W. J. (1993). E-optimal designs for polynomial regression. *Ann. Statist.* **21** 402–415. MR1212184

RAFAJLOWICZ, E. and MYSZKA, W. (1988). Optimum experimental design for a regression on a hypercube-generalization of Hoel's result. *Ann. Inst. Statist. Math.* **40** 821–827. MR0996701

SILVEY, S. D. (1980). *Optimal Design*. Chapman & Hall, London. MR0606742

FAKULTÄT FÜR MATHEMATIK
RUHR-UNIVERSITÄT BOCHUM
44780 BOCHUM
GERMANY
E-MAIL: holger.dette@rub.de

DEPARTMENT OF COMPUTER SCIENCE
ST.-PETERSBURG STATE ELECTROTECHNICAL UNIVERSITY
197376 ST. PETERSBURG
RUSSIA