Statistica Sinica **19** (2009), 1055-1075

MODEL-ROBUST D- AND A-OPTIMAL DESIGNS FOR MIXTURE EXPERIMENTS

Mong-Na Lo Huang[1], Hsiang-Ling Hsu[1], Chao-Jin Chou[1] and Thomas Klein[2]

[1]*National Sun Yat-sen University and* [2]*Technische Universität München*

Abstract: This paper investigates optimal designs for mixture experiments when there is uncertainty as to whether a polynomial regression model of degree one or two is appropriate. Three groups of novel results are presented: (i) a complete class of designs relative to certain mixed design criteria, (ii) model-robust D- and A-optimal designs, (iii) D- and A-optimal designs with maximin efficiencies under variation of the design criterion.

Key words and phrases: centroid design, complete class, design efficiency, design optimality, exchangeability, Kiefer ordering, polynomial regression, robustness.

1. Introduction

A *mixture experiment* is an experiment in which the factors $x_1, \ldots, x_q$ $(q \geq 2)$ are non-negative and sum to unity, that is, the factors represent relative proportions of the q ingredients blended in a mixture. The experimental conditions are thus elements of the *probability simplex* $\mathcal{S}^{q-1} := \{\mathbf{x} \in [0,1]^q : \mathbf{x}'1_q = 1\}$, with $1_q := (1, \ldots, 1)' \in \mathbb{R}^q$. Cornell (2002) has numerous examples and applications of mixture experiments.

Various types of models have been proposed for mixture experiments, such as polynomial models of degree $d \leq 3$, models with inverse terms, and log-contrast models. In this paper we investigate mixture experiments where there is uncertainty as to whether Scheffé's (1958) first- or second-degree polynomial mixture model is appropriate. Uncertainty about the choice of model is frequently encountered in practical applications and should be taken into account at the design stage. The *regression functions* specifying the first- and second-degree polynomial mixture models are

$$f_1 : \mathcal{S}^{q-1} \to \mathbb{R}^{m_1},\ \mathbf{x} \mapsto \mathbf{x},$$
$$f_2 : \mathcal{S}^{q-1} \to \mathbb{R}^{m_2},\ \mathbf{x} = (x_1, \ldots, x_q)' \mapsto \big(\mathbf{x}', (x_i x_j)_{1 \leq i < j \leq q}\big)',$$

with $m_1 := q$ and $m_2 := \binom{q+1}{2}$. We refer to these models as $(\mathcal{S}^{q-1}, f_1)$ and $(\mathcal{S}^{q-1}, f_2)$, and they are written as

$$\mathrm{E}[y(\mathbf{x})] = \theta' f_1(\mathbf{x}) = \sum_{i=1}^{q} \theta_i x_i\,, \qquad \mathrm{E}[y(\mathbf{x})] = \beta' f_2(\mathbf{x}) = \sum_{i=1}^{q} \beta_i x_i + \sum_{\substack{i,\,j=1\\ i<j}}^{q} \beta_{ij} x_i x_j,$$

with unknown parameter vectors $\theta \in \mathbb{R}^{m_1}$ and $\beta \in \mathbb{R}^{m_2}$, respectively. All observations taken in an experiment are assumed to be uncorrelated and to have common unknown variance $\sigma^2 > 0$. Note that $(\mathcal{S}^{q-1}, f_1)$ is a proper submodel of $(\mathcal{S}^{q-1}, f_2)$ since

$$f_1(\mathbf{x}) = Z f_2(\mathbf{x}) \qquad \text{for all } \mathbf{x} \in \mathcal{S}^{q-1}, \tag{1.1}$$

with $Z := (I_{m_1}, 0_{m_1 \times (m_2 - m_1)}) \in \mathbb{R}^{m_1 \times m_2}$, and I_{m_1} denoting the $m_1 \times m_1$ identity matrix. An *experimental design* ξ for a mixture model is a probability measure on $\mathcal{S}^{q-1}$ with finite support, $|\operatorname{supp} \xi| < \infty$. The statistical properties of a design ξ within the model $(\mathcal{S}^{q-1}, f_i)$ are captured by its *information matrix* $M_i(\xi) := \int_{\mathcal{S}^{q-1}} f_i(\mathbf{x}) f_i'(\mathbf{x})\, \mathrm{d}\xi(\mathbf{x})$ for $i = 1, 2$. Note that (1.1) implies

$$M_1(\xi) = Z M_2(\xi) Z' \qquad \text{for all designs } \xi \text{ on } \mathcal{S}^{q-1}. \tag{1.2}$$

A design ξ is called *feasible* in a given model if and only if its information matrix has full rank. Associated with any feasible design in the model $(\mathcal{S}^{q-1}, f_i)$ we have the *dispersion function* $d_i(., \xi) : \mathcal{S}^{q-1} \to [0, \infty)$, $\mathbf{x} \mapsto f_i'(\mathbf{x}) M_i^{-1}(\xi) f_i(\mathbf{x})$, for $i = 1, 2$.

Let $\mathrm{PD}(m_i)$ denote the cone of positive definite $m_i \times m_i$ matrices, for $i = 1, 2$. Given $(\mathcal{S}^{q-1}, f_i)$, $i = 1, 2$, we measure a feasible design's information matrix using the scalar functions $\phi_i^D, \phi_i^A : \mathrm{PD}(m_i) \to (0, \infty)$ defined by

$$\phi_i^D(M) := |M| = \det M \qquad \text{and} \qquad \phi_i^A(M) := (\operatorname{tr} M^{-1})^{-1},$$

and called the D- and the *A-criterion*, respectively. Both the D- and the A-criterion are members of the popular family of Kiefer's ϕ_p-criteria, see Pukelsheim (2006). A feasible design ξ^* is called *D-optimal* (or *A-optimal*) within the model $(\mathcal{S}^{q-1}, f_i)$ if it maximizes $\phi_i^D\big(M_i(\xi)\big)$ (or $\phi_i^A\big(M_i(\xi)\big)$, respectively) among all feasible designs ξ on $\mathcal{S}^{q-1}$. Necessary and sufficient conditions for optimality are given in Pukelsheim (2006). For instance, a feasible design ξ^* is D-optimal in model $(\mathcal{S}^{q-1}, f_i)$ if and only if it minimizes $\max\{d_i(\mathbf{x}, \xi) \,|\, \mathbf{x} \in \mathcal{S}^{q-1}\}$ among all feasible designs ξ on $\mathcal{S}^{q-1}$. In this case we have $\max\{d_i(\mathbf{x}, \xi^*) \,|\, \mathbf{x} \in \mathcal{S}^{q-1}\} = q$, where the maximum is attained at all points $\mathbf{x} \in \operatorname{supp} \xi^*$. Similarly, a design ξ^* is A-optimal in model $(\mathcal{S}^{q-1}, f_i)$ if and only if $f_i'(\mathbf{x}) M_i^{-2}(\xi^*) f_i(\mathbf{x}) \le \operatorname{tr} M_i^{-1}(\xi^*)$ for all $\mathbf{x} \in \mathcal{S}^{q-1}$.

Designs for mixture experiments have been investigated extensively in the literature, see Chan's (2000) comprehensive overview. Kiefer (1961) derived D-optimal designs in Scheffé's second-degree model. Galil and Kiefer (1977) presented numerical results on ϕ_p-optimal designs in that model. Our starting point is the completeness results of Draper and Pukelsheim (1999) and Draper, Heiligers and Pukelsheim (2000) that reduce the optimal design problem to a mere allocation problem. We make particular use of Klein's (2004b) method for evaluating Kiefer-Wolfowitz-type optimality conditions based on the analysis of a quadratic subspace of invariant symmetric matrices, see Klein (2004a). Concern about model uncertainty dates back to Box and Draper (1959), and various design strategies addressing this issue have been suggested, see Atkinson and Cox (1974), Studden (1982), Huang and Studden (1988), Dette (1990, 1991, 1993), and Pukelsheim and Rosenberger (1993). Our design approach toward model uncertainty can also be viewed as a compound or multi-purpose design strategy, see Läuter (1974, 1976) or Cook and Wong (1994).

The paper is organized as follows. Section 2 introduces our models' symmetry structure and gives a completeness result for the class of weighted centroid designs. Due to this result, we can reduce the set of competitors in Sections 3 and 4, where two model-robust D- and A-criteria are defined and model-robust D- and A-optimal designs are derived. In Section 5, we explore a maximin approach for the efficiencies of model-robust D- and A- optimal designs under variation of the optimality criterion. A final section summarizes the paper.

2. Kiefer Orderings and Mixed Design Criteria

We start by discussing the invariance properties of the design problems investigated in Sections 3−5, very much in the spirit of Kiefer (1961) and Pukelsheim ((2006), Chap. 13, 14). Denote the canonical unit vectors in $\mathbb{R}^q$ by $e_1, \ldots, e_q$, and those in $\mathbb{R}^{\binom{q}{2}}$ by e_{ij} with $1 \le i < j \le q$, and let $\mathfrak{S}_q$ denote the group of permutations of $\{1, \ldots, q\}$. The probability simplex $\mathcal{S}^{q-1}$ has a natural symmetry structure, it is invariant under permutation of the coordinates. Formally we have $R_\pi \mathcal{S}^{q-1} = \mathcal{S}^{q-1}$ for all $R_\pi := \sum_{i=1}^q e_{\pi(i)} e_i'$, $\pi \in \mathfrak{S}_q$, that is, the group $\mathrm{Perm}(q)$ of $q \times q$ permutation matrices acts on $\mathcal{S}^{q-1}$ through $(R, \mathbf{x}) \mapsto R\mathbf{x}$. This induces an action $(R, \xi) \mapsto \xi^R = \xi \circ R^{-1}$ of $\mathrm{Perm}(q)$ on the set of all designs on $\mathcal{S}^{q-1}$. A design ξ satisfying $\xi = \xi^R$ for all $R \in \mathrm{Perm}(q)$ is called *exchangeable*. Now we define the groups $\mathcal{H}_1 := \mathrm{Perm}(m_1)$ and

$$\mathcal{H}_2 := \left\{ H_\pi = \begin{pmatrix} R_\pi 0 \\ 0 S_\pi \end{pmatrix} : \pi \in \mathfrak{S}_q \right\}, \quad S_\pi := \sum_{\substack{i,j=1 \\ i<j}}^{q} e_{(\pi(i),\pi(j))_\uparrow} \, e_{ij}' \,, \; \pi \in \mathfrak{S}_q,$$

where $(\pi(i), \pi(j))_\uparrow$ denotes the pair of indices $\pi(i)$, $\pi(j)$ in ascending order. Note that $\mathcal{H}_2$ is a subgroup of $\mathrm{Perm}(m_2)$, see Klein (2004a). The group $\mathcal{H}_i$ acts on the cone $\mathrm{NND}(m_i)$ of nonnegative definite $m_i \times m_i$ matrices by

$$\mathcal{H}_i \times \mathrm{NND}(m_i) \to \mathrm{NND}(m_i), \quad (H, M) \mapsto HMH', \qquad \text{for } i = 1, 2. \tag{1.3}$$

These actions are linked to that of $\mathrm{Perm}(q)$ on $\mathcal{S}^{q-1}$ by the regression functions' *equivariance properties*, $f_1(R_\pi \mathbf{x}) = R_\pi f_1(\mathbf{x})$ and $f_2(R_\pi \mathbf{x}) = H_\pi f_2(\mathbf{x})$ for all

$\mathbf{x} \in \mathcal{S}^{q-1}$, $\pi \in \mathfrak{S}_q$. For any design ξ these imply $M_1(\xi^{R_\pi}) = R_\pi M_1(\xi) R'_\pi$ and $M_2(\xi^{R_\pi}) = H_\pi M_2(\xi) H'_\pi$ for all $\pi \in \mathfrak{S}_q$. Consequently, the information matrices of an exchangeable design ξ are $\mathcal{H}_i$-invariant or *exchangeable*,

$$M_i(\xi) = H M_i(\xi) H' \qquad \text{for all } H \in \mathcal{H}_i,\ i = 1, 2. \tag{1.4}$$

Each of the group actions from (1.3) induces a *Kiefer ordering* of information matrices, see Pukelsheim (2006). For $i = 1, 2$, and $A, B \in \mathrm{NND}(m_i)$, we write $A \ll_{\mathcal{H}_i} B$ (B improves upon A in the Kiefer ordering induced by $\mathcal{H}_i$) if there is a matrix $C \in \mathrm{conv}\,\{HAH' \mid H \in \mathcal{H}_i\}$ such that $C \le B$. Here $\le$ denotes the *Löwner ordering* on $\mathrm{NND}(m_i)$, defined by $C \le B$ if and only if $B - C \in \mathrm{NND}(m_i)$. The Kiefer ordering's significance lies in the fact that any design criterion that is Löwner monotonic, matrix-concave, and invariant under the action of $\mathcal{H}_i$ on $\mathrm{NND}(m_i)$, is monotonic with respect to $\ll_{\mathcal{H}_i}$. In particular, this is the case for the orthogonally invariant criteria ϕ_i^D and ϕ_i^A. The following lemma establishes a relation between $\ll_{\mathcal{H}_1}$ and $\ll_{\mathcal{H}_2}$.

Lemma 2.1. *Let ξ, ξ' be two designs on $\mathcal{S}^{q-1}$. Then the inequality $M_2(\xi) \ll_{\mathcal{H}_2} M_2(\xi')$ implies $M_1(\xi) \ll_{\mathcal{H}_1} M_1(\xi')$.*

Proof. First we observe $ZH_\pi = R_\pi Z$ for all $\pi \in \mathfrak{S}_q$, with Z from (1.1). Now assume $M_2(\xi) \ll_{\mathcal{H}_2} M_2(\xi')$, that is, $M_2(\xi') \ge C$ for some $C \in \mathrm{conv}\,\{HM_2(\xi)H' \mid H \in \mathcal{H}_2\}$. Writing $C = \sum_{\pi \in \mathfrak{S}_q} \gamma_\pi H_\pi M_2(\xi) H'_\pi$ with $\gamma_\pi \ge 0$, $\pi \in \mathfrak{S}_q$, $\sum_{\pi \in \mathfrak{S}_q} \gamma_\pi = 1$, we find

$$ZCZ' = \sum_{\pi \in \mathfrak{S}_q} \gamma_\pi Z H_\pi M_2(\xi) H'_\pi Z' = \sum_{\pi \in \mathfrak{S}_q} \gamma_\pi R_\pi Z M_2(\xi) Z' R'_\pi = \sum_{\pi \in \mathfrak{S}_q} \gamma_\pi R_\pi M_1(\xi) R'_\pi,$$

which proves $ZCZ' \in \mathrm{conv}\,\{HM_1(\xi)H' \mid H \in \mathcal{H}_1\}$. Our assumption $C \le M_2(\xi')$ implies $ZCZ' \le ZM_2(\xi')Z' = M_1(\xi')$, thus establishing the assertion.

In Sections 3 and 4 we explore certain *mixed design criteria*, that is, functions of a design criterion in the first-degree model and another criterion in the second-degree model. Now we link such criteria to the Kiefer ordering $\ll_{\mathcal{H}_2}$ using Lemma 2.1.

Lemma 2.2. *Let $\phi_i : \mathrm{NND}(m_i) \to [0, \infty)$, $i = 1, 2$, be functions that are monotonic relative to the Kiefer orderings $\ll_{\mathcal{H}_i}$, respectively. Furthermore, let $\varphi : [0, \infty)^2 \to [0, \infty)$ be a function that is increasing in both arguments. Finally, define $\Phi : \Xi \to [0, \infty)$, $\xi \mapsto \varphi\big(\phi_1(M_1(\xi)), \phi_2(M_2(\xi))\big)$, where Ξ denotes the set of all designs on $\mathcal{S}^{q-1}$. Then the inequality $M_2(\xi) \ll_{\mathcal{H}_2} M_2(\xi')$ implies $\Phi(\xi) \le \Phi(\xi')$, for all $\xi, \xi' \in \Xi$.*

Proof. Assume $M_2(\xi) \ll_{\mathcal{H}_2} M_2(\xi')$ with $\xi, \xi' \in \Xi$. By Lemma 2.1 we also have $M_1(\xi) \ll_{\mathcal{H}_1} M_1(\xi')$. The claim follows from the monotonicity of ϕ_1, ϕ_2, and φ.

The above Lemma 2.2 allows us to restrict our considerations to the class of *weighted centroid designs* defined below. For $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}^{q-1}$ we use the notation $\mathbf{x}_1 \leftrightarrow \mathbf{x}_2$ to indicate $\mathbf{x}_1 = R\mathbf{x}_2$ for some $R \in \mathrm{Perm}(q)$.

Definition 2.3. For $1 \le j \le q$, the *j-th elementary centroid design* η_j is the uniform distribution on the $\binom{q}{j}$ points of the form $\mathbf{x} \leftrightarrow j^{-1}\sum_{k=1}^{j} e_k \in \mathcal{S}^{q-1}$. A convex combination $\eta(\lambda) = \sum_{j=1}^{q} \lambda_j \eta_j$ with $\lambda = (\lambda_1, \ldots, \lambda_q)' \in \mathcal{S}^{q-1}$ is called a *weighted centroid design* with *weight vector* λ. We write $\mathcal{W}$ for the set of all weighted centroid designs.

Theorem 2.4. *Let ϕ_1, ϕ_2, φ, and Φ satisfy the assumptions of Lemma* 2.2. *Then the class $\mathcal{W}$ of weighted centroid designs is essentially complete relative to Φ, that is, for every design ξ on $\mathcal{S}^{q-1}$, there is $\eta \in \mathcal{W}$ such that $\Phi(\xi) \le \Phi(\eta)$.*

Proof. Draper and Pukelsheim (1999) and Draper et al. (2000) proved the class $\mathcal{W}$ essentially complete relative to $\ll_{\mathcal{H}_2}$, that is, for every design ξ on $\mathcal{S}^{q-1}$ there is $\eta \in \mathcal{W}$ such that $M_2(\xi) \ll_{\mathcal{H}_2} M_2(\eta)$. Hence the claim follows from Lemma 2.2.

Weighted centroid designs are exchangeable by definition. Their information matrices in the model $(\mathcal{S}^{q-1}, f_i)$, $i = 1, 2$, are thus $\mathcal{H}_i$-invariant, see (1.4). In particular, they are elements of $\mathrm{Sym}(m_i, \mathcal{H}_i) := \{M \in \mathrm{Sym}(m_i) \mid HMH' = M \text{ for all } H \in \mathcal{H}_i\}$, the subspace of *$\mathcal{H}_i$-invariant symmetric matrices*, where Sym (m_i) denotes the space of symmetric $m_i \times m_i$ matrices. The following lemma quotes a fundamental result on the matrix subspaces $\mathrm{Sym}(m_i, \mathcal{H}_i)$, $i = 1, 2$, from Klein (2004a).

Lemma 2.5. *Let $U_1 := I_q$, $U_2 := 1_q 1_q' - I_q$, $W_1 := I_{\binom{q}{2}}$, and*

$$V_1 := \sum_{\substack{i,j=1 \\ i<j}}^{q} e_{ij}\,(e_i + e_j)', \qquad V_2 := \sum_{\substack{i,j=1 \\ i<j}}^{q} \sum_{\substack{k=1 \\ k \notin \{i,j\}}}^{q} e_{ij}\, e_k',$$

$$W_2 := \sum_{\substack{i,j=1 \\ i<j}}^{q} \sum_{\substack{k,\ell=1 \\ k<\ell \\ |\{i,j\} \cap \{k,\ell\}| = 1}}^{q} e_{ij}\, e_{k\ell}', \qquad W_3 := \sum_{\substack{i,j=1 \\ i<j}}^{q} \sum_{\substack{k,\ell=1 \\ k<\ell \\ \{i,j\} \cap \{k,\ell\} = \varnothing}}^{q} e_{ij}\, e_{k\ell}'.$$

Then any matrix $M \in \mathrm{Sym}(m_1, \mathcal{H}_1)$ can be uniquely represented as $M = a\,U_1 + b\,U_2$ with coefficients $a, b \in \mathbb{R}$. Similarly, any matrix $M \in \mathrm{Sym}(m_2, \mathcal{H}_2)$ is of the form

$$M = \begin{pmatrix} a\,U_1 + b\,U_2 & \textit{symm.} \\ c\,V_1 + d\,V_2 & e\,W_1 + f\,W_2 + g\,W_3 \end{pmatrix}$$

with unique coefficients $a, \ldots, g \in \mathbb{R}$. Note that $V_2 = 0$, $W_2 = W_3 = 0$ for $q = 2$, and $W_3 = 0$ for $q = 3$.

Due to the fact that $\mathcal{H}_i$, $i = 1, 2$, is a subgroup of the orthogonal group, the subspace $\mathrm{Sym}(m_i, \mathcal{H}_i)$ is a *quadratic subspace*, that is, it is closed under formation of powers $M \mapsto M^n$ for $n \in \mathbb{Z}$, see Pukelsheim ((2006), p. 346). Computing such powers is a crucial step in evaluating optimality conditions related to the D- and the A-criterion. We make extensive use of a multiplication table for the matrices $U_1, \ldots, W_3$ given by Klein ((2004a), Lemma 3.2) that allows efficient computation of these matrix powers.

3. Model-Robust D-Optimal Designs

The unique D-optimal design in Scheffé's first-degree model is $\xi_1^D := \eta_1$, see Draper and Pukelsheim (1999). In the second-degree model, the unique D-optimal design is $\xi_2^D := [2/(q+1)]\eta_1 + [(q-1)/(q+1)]\eta_2$, see Kiefer (1961). If the experimenter is uncertain about the appropriate model, then neither of the two results is particularly useful. For instance, the design ξ_2^D is non-optimal in the first-degree model, and the design ξ_1^D is not even feasible in the second-degree model. This raises the central question of the present paper: Which designs shall an experimenter use if there is uncertainty about whether the first- or the second-degree model is more appropriate for a given mixture experiment? We answer this question by assessing designs in terms of a *model-robust D-criterion.*

Our criterion is a tradeoff between the D-criteria in the first- and second-degree models. For $r \in [0, 1]$ the model-robust D-criterion is defined as a weighted geometric average of the D-efficiencies of a design ξ under the first- and second-degree models $\big(|M_1(\xi)|/|M_1(\xi_1^D)|\big)^{r/m_1} \big(|M_2(\xi)|/|M_2(\xi_2^D)|\big)^{(1-r)/m_2}$; here $|M_1(\xi_1^D)|$ and $|M_2(\xi_2^D)|$ are fixed constants. Maximizing this criterion is equivalent to maximizing $\psi_r(\xi) := |M_1(\xi)|^{r/m_1} \, |M_2(\xi)|^{(1-r)/m_2}$, a weighted average of the two D-optimality criteria. The parameter r is a prior, that is, an initial weight assigned to the first-degree model. Taking logarithms we obtain the *model-robust D-criterion*

$$\Psi_r^D(\xi) := \log \psi_r(\xi) = \frac{r}{m_1} \log(|M_1(\xi)|) + \frac{1-r}{m_2} \log(|M_2(\xi)|). \tag{3.1}$$

A design ξ_r^D with $M_2(\xi_r^D) \in \mathrm{PD}(m_2)$ is called *model-robust D-optimal* if and only if it satisfies

$$\Psi_r^D(\xi_r^D) = \max\{\Psi_r^D(\xi) \mid \xi \in \Xi \text{ with } M_2(\xi) \in \mathrm{PD}(m_2)\}. \tag{3.2}$$

Note that $M_2(\xi) \in \mathrm{PD}(m_2)$ implies $M_1(\xi) \in \mathrm{PD}(m_1)$. As in Section 1, we have an associated dispersion function

$$d(.,\xi) : \mathcal{S}^{q-1} \to [0, \infty], \quad \mathbf{x} \mapsto \frac{r}{m_1} d_1(\mathbf{x}, \xi) + \frac{1-r}{m_2} d_2(\mathbf{x}, \xi) \tag{3.3}$$

for every design ξ on $\mathcal{S}^{q-1}$. In a similar fashion to Dette's (1990) Equivalence Theorem, it can be shown that a design ξ_r^D with $M_2(\xi) \in \mathrm{PD}(m_2)$ is model-robust D-optimal for a given prior $r \in [0,1]$ if and only if its dispersion function satisfies

$$d(\mathbf{x}, \xi_r^D) \le 1 \quad \text{for all } \mathbf{x} \in \mathcal{S}^{q-1}. \tag{3.4}$$

As a first step toward solving the design problem (3.2) we restrict the set of competing designs to the class of weighted centroid designs.

Lemma 3.1. *The set $\mathcal{W}$ from Definition* 2.3 *is an essentially complete class of designs relative to the model-robust D-criterion Ψ_r^D, $r \in [0,1]$, defined in* (3.1). *Then a design ξ_r^D with $M_2(\xi_r^D) \in \mathrm{PD}(m_2)$ (for given $r \in [0,1]$) is model-robust D-optimal if and only if $d(\mathbf{x}, \xi_r^D) \le 1$ for all $\mathbf{x} \in \eta_j$, $1 \le j \le q$.*

Proof. Upon setting $\varphi(y_1, y_2) := (r/m_1)\log y_1 + [(1-r)/m_2]\log y_2$ we can write $\Psi_r^D(\xi) = \varphi\big(\phi_1^D(M_1(\xi)), \phi_2^D(M_2(\xi))\big)$ for all designs ξ. Clearly, ϕ_1^D, ϕ_2^D and ϕ satisfy the assumptions of Lemma 2.2, which is why Theorem 2.3 establishes the claim. The set of competitors in the design problem (3.2) can thus be restricted to $\mathcal{W}$. Based on standard arguments the optimality condition (3.4) then simplifies to the one stated in the second claim.

Note that the above proof yields an even stronger result than the one stated in Lemma 3.1: for any design ξ on $\mathcal{S}^{q-1}$ there is $\eta \in \mathcal{W}$ such that $\Psi_r^D(\xi) \le \Psi_r^D(\eta)$ for all $r \in [0,1]$. In this sense the completeness property of $\mathcal{W}$ holds uniformly with respect to the family of criteria $\big(\Psi_r^D \big| r \in [0,1]\big)$.

The optimality condition in Lemma 3.1 directs us toward evaluating the dispersion function of a candidate design in support points of weighted centroid designs.

Lemma 3.2. *For every exchangeable design ξ on $\mathcal{S}^{q-1}$, the dispersion function given in* (3.3) *satisfies*

$$d(\mathbf{x}, \xi) = \frac{r}{m_1}\,\mathrm{tr}\,(M_1(\eta_j)\, M_1^{-1}(\xi)) + \frac{1-r}{m_2}\,\mathrm{tr}\,(M_2(\eta_j)\, M_2^{-1}(\xi))$$

for all $\mathbf{x} \in \operatorname{supp}\eta_j$, $1 \le j \le q$.

Proof. Assume $\mathbf{x} \in \mathcal{S}^{q-1}$, and let $\varepsilon_{\mathbf{x}}$ denote the single-point design in $\mathbf{x}$. Then the balanced design $\bar{\varepsilon}_{\mathbf{x}}$ puts equal weights on all points $\mathbf{y} \in \mathcal{S}^{q-1}$ with $\mathbf{y} \leftrightarrow \mathbf{x}$. Standard arguments based on $\mathcal{H}_i$-invariance of information matrices show $d_i(\mathbf{x}, \xi) = \mathrm{tr}\,\big(M_i(\bar{\varepsilon}_{\mathbf{x}})\, M_i^{-1}(\xi)\big)$ for all exchangeable designs ξ and $i = 1, 2$. From the definition of $d(., \xi)$, we obtain

$$d(\mathbf{x}, \xi) = \frac{r}{m_1}\,\mathrm{tr}\,\big(M_1(\bar{\varepsilon}_{\mathbf{x}}) M_1^{-1}(\xi)\big) + \frac{1-r}{m_2}\,\mathrm{tr}\,\big(M_2(\bar{\varepsilon}_{\mathbf{x}}) M_2^{-1}(\xi)\big).$$

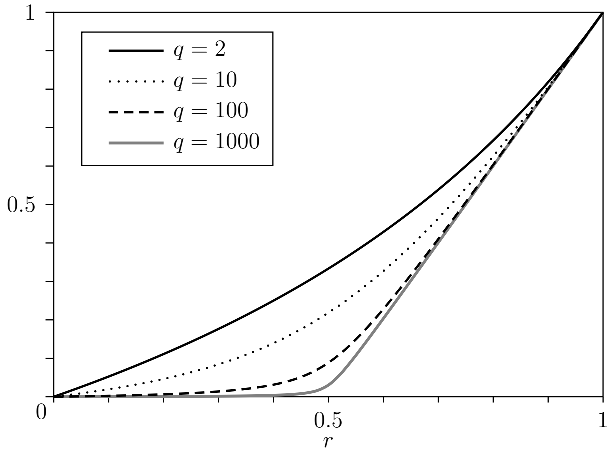

Figure 1. The weight α_r^D as a function of $r \in [0,1]$

Noting that $\bar{\varepsilon}_{\mathbf{x}} = \eta_j$ if and only if $\mathbf{x} \in \operatorname{supp} \eta_j$, the assertion is proved.

The following definition is essential to our approach of determining model-robust D-optimal designs.

Definition 3.3. For all $\alpha \in [0,1]$ we write $\xi_\alpha := \alpha\xi_1^D + (1-\alpha)\xi_2^D$, where ξ_i^D is the D-optimal design in the model $(\mathcal{S}^{q-1}, f_i)$, $i = 1, 2$. In addition, we define the set of designs $\Xi_D := \operatorname{conv}\{\xi_1^D, \xi_2^D\} = \{\xi_\alpha | \, \alpha \in [0,1]\} \subseteq \mathcal{W}$.

Now we solve the problem of maximizing the model-robust D-criterion (3.1) in two steps. First, we generate an optimality candidate by finding a weight $\alpha_r^D \in [0,1]$ such that $\xi_{\alpha_r^D} \in \Xi_D$ maximizes (3.1) within Ξ_D. Second, we prove the candidate design $\xi_{\alpha_r^D}$ to be optimal among *all* feasible designs. The following lemma produces the optimality candidate $\xi_{\alpha_r^D}$.

Lemma 3.4. *For any $r \in [0,1]$, the design $\xi_{\alpha_r^D} \in \Xi_D$ with weight*

$$\alpha_r^D = \frac{q\,(2r-1) - 2 - r + \sqrt{8r(q-r) + (2+q+r-2qr)^2}}{2(q-r)} \in (0,1)$$

is the unique model-robust D-optimal design among all designs in Ξ_D. (Figure 1 displays the function $r \mapsto \alpha_r^D$ for various choices of q.)

Proof. For all $\alpha \in (0,1)$ we have

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\Psi_r^D(\xi_\alpha) = \frac{r}{m_1}\operatorname{tr} M_1^{-1}(\xi_\alpha)[M_1(\xi_1^D) - M_1(\xi_2^D)] + \frac{1-r}{m_2}\operatorname{tr} M_2^{-1}(\xi_\alpha)[M_2(\xi_1^D) - M_2(\xi_2^D)]. \quad (3.5)$$

Adopting the block-matrix notation from Klein (2004a), we can write $M_1(\xi_1^D) = U_1/q$, $M_1(\xi_2^D) = [(q+3)U_1 + U_2]/(4\binom{q+1}{2})$, and

$$M_2(\xi_1^D) = \begin{pmatrix} M_1(\xi_1^D) & 0 \\ 0 & 0 \end{pmatrix}, \qquad M_2(\xi_2^D) = \begin{pmatrix} M_1(\xi_2^D) & \frac{1}{8\binom{q+1}{2}} V_1' \\ \frac{1}{8\binom{q+1}{2}} V_1 & \frac{1}{16\binom{q+1}{2}} W_1 \end{pmatrix}.$$

From $M_i(\xi_\alpha) = \alpha M_i(\xi_1^D) + (1-\alpha) M_i(\xi_2^D)$ we obtain

$$M_1(\xi_\alpha) = \frac{1}{2q(q+1)} \big([(q-1)\alpha + q + 3]\, U_1 + (1-\alpha)\, U_2 \big),$$
$$M_2(\xi_\alpha) = \begin{pmatrix} M_1(\xi_\alpha) & \frac{1-\alpha}{4q(q+1)} V_1' \\ \frac{1-\alpha}{4q(q+1)} V_1 & \frac{1-\alpha}{8q(q+1)} W_1 \end{pmatrix}.$$

Using the inverse formula in Fedorov ((1972), pp. 16, 17) and the multiplication table from Lemma 3.2 in Klein (2004a), we compute the inverse information matrices

$$M_1^{-1}(\xi_\alpha) = \bar{a}_1 U_1 + \bar{b}_1 U_2 \qquad \text{and} \qquad M_2^{-1}(\xi_\alpha) = \begin{pmatrix} \bar{a}_2 U_1 & \bar{c}_2 V_1' \\ \bar{c}_2 V_1 & \bar{e}_2 W_1 + \bar{f}_2 W_2 \end{pmatrix} \tag{3.6}$$

with coefficients

$$\bar{a}_1 = \frac{q(2q+1+\alpha)}{(q+2)+q\alpha}, \qquad \bar{b}_1 = -\frac{q(1-\alpha)}{(q+2)+q\alpha},$$
$$\bar{a}_2 = \frac{q(q+1)}{2+(q-1)\alpha}, \qquad \bar{c}_2 = -\frac{2q(q+1)}{2+(q-1)\alpha},$$
$$\bar{e}_2 = \frac{8q(q+1)[3+(q-2)\alpha]}{(1-\alpha)[2+(q-1)\alpha]}, \qquad \bar{f}_2 = \frac{4q(q+1)}{2+(q-1)\alpha}.$$

Further application of the abovementioned multiplication table and $\operatorname{tr} U_2 = \operatorname{tr} W_2 = \operatorname{tr} W_3 = 0$ yield

$$\operatorname{tr} M_1^{-1}(\xi_\alpha) M_1(\xi_1^D) = \bar{a}_1, \tag{3.7}$$

$$\operatorname{tr} M_1^{-1}(\xi_\alpha) M_1(\xi_2^D) = \frac{2}{q+1} \left(a_1 + \frac{(q-1)(\bar{a}_1 + \bar{b}_1)}{4} \right), \tag{3.8}$$

$$\operatorname{tr} M_2^{-1}(\xi_\alpha) M_2(\xi_1^D) = \bar{a}_2, \tag{3.9}$$

$$\operatorname{tr} M_2^{-1}(\xi_\alpha) M_2(\xi_2^D) = \frac{q}{q+1} \left(\frac{(q+3)\bar{a}_2}{2q} + \frac{(q-1)\bar{c}_2}{4(q+2)} \right) + \frac{q(q-1)}{4(q+1)} \left(\frac{\bar{c}_2}{2q} + \frac{\bar{e}_2}{8(q+2)} \right). \tag{3.10}$$

Substituting (3.7)–(3.10) into (3.5), we see that $\frac{\mathrm{d}}{\mathrm{d}\alpha} \Psi_r^D(\xi_\alpha) = 0$ holds if and only if

$$r = \frac{(q+2)\alpha_r^D + q(\alpha_r^D)^2}{2 + (2q-1)\alpha_r^D + (\alpha_r^D)^2}.$$

The unique solution of this equation in $[0,1]$ is the weight α_r^D given in the claim.

Using the Equivalence Theorem from Lemma 3.1, we now prove the candidate designs $\xi_{\alpha_r^D}$ to be model-robust D-optimal among *all* designs. We follow Klein's (2004b) approach which is based on the quadratic subspace properties introduced in Section 2 and yields tractable matrix algebra.

Theorem 3.5. *The design $\xi_{\alpha_r^D} \in \Xi_D$ given in Lemma 3.4 is a model-robust D-optimal design for Scheffé's first-degree and second-degree models.*

Proof. We check the optimality condition provided by Lemma 3.1. Assume $j \in \{1,\ldots,q\}$ and $\mathbf{x} \in \operatorname{supp} \eta_j$. Since the design $\xi_{\alpha_r^D}$ is exchangeable, we can apply Lemma 3.2 and obtain

$$d(\mathbf{x}, \xi_{\alpha_r^D}) = \frac{r}{m_1} \operatorname{tr}(M_1^{-1}(\xi_{\alpha_r^D})\, M_1(\eta_j)) + \frac{1-r}{m_2} \operatorname{tr}(M_2^{-1}(\xi_{\alpha_r^D})\, M_2(\eta_j)) =: N_j\,.$$

Based on the representation of $\mathcal{H}_i$-invariant symmetric matrices ($i = 1, 2$) from Lemma 2.5, and the moments of elementary centroid designs given by Draper et al. (2000), the involved information matrices can be written as

$$M_1(\eta_j) = a_j U_1 + b_j U_2 \quad \text{and} \quad M_2(\eta_j) = \begin{pmatrix} M_1(\eta_j) & c_j V_1' + d_j V_2' \\ c_j V_1 + d_j V_2 & e_j W_1 + f_j W_2 + g_j W_3 \end{pmatrix}$$

with coefficients

$$\begin{aligned} a_j &= \frac{1}{jq}\,, & b_j &= \frac{j-1}{jq(q-1)}\,, \\ c_j &= \frac{j-1}{j^2 q(q-1)}\,, & d_j &= \frac{(j-1)(j-2)}{j^2 q(q-1)(q-2)}\,, \\ e_j &= \frac{j-1}{j^3 q(q-1)}\,, & f_j &= \frac{(j-1)(j-2)}{j^3 q(q-1)(q-2)}\,, & g_j &= \frac{(j-1)(j-2)(j-3)}{j^3 q(q-1)(q-2)(q-3)}\,. \end{aligned}$$

With the inverse matrices found in (3.6) and the multiplication table from Lemma 3.2 in Klein (2004a), lengthy calculations yield

$$N_j = \frac{2\,(j^2 + 4j - 4)\,(1 + \alpha_r^D q) - j^3\, \alpha_r^D\, (1 - \alpha_r^D)}{j^3\,[\,(\alpha_r^D)^2 + (2q-1)\alpha_r^D + 2\,]} \qquad \text{for all } 1 \le j \le q.$$

Computing differences of these terms leads to

$$N_j - N_{j+1} = \frac{2\,[-4 - 8j + 8j^3 + j^2\,(1+j)^2]\,(1 + \alpha_r^D q)}{j^3\,(j+1)^3\,[(\alpha_r^D)^2 + (2q-1)\alpha_r^D + 2]} \qquad \text{for all } 1 \le j \le q-1.$$

Recalling $q \ge 2$, we see $N_1 = N_2 = 1$ and $N_j - N_{j+1} > 0$ for $j \ge 2$, that is, $N_1 = N_2 > \cdots > N_q$. This sequence of inequalities shows

$$d(\mathbf{x}, \xi_{\alpha_r^D}) \begin{cases} = 1 \text{ for all } \mathbf{x} \in (\operatorname{supp} \eta_1 \cup \operatorname{supp} \eta_2)\,, \\ < 1 \text{ for all } \mathbf{x} \in \bigcup_{j=3}^{q} \operatorname{supp} \eta_j\,. \end{cases}$$

Thus the equivalence theorem is satisfied and the proof is completed.

4. Model-Robust A-Optimal Designs

We now adapt the arguments used in Section 3 to the alternative concept of *model-robust A-optimality.* The D-optimal design ξ_1^D in Scheffé's first-degree model is also A-optimal, $\xi_1^A := \xi_1^D = \eta_1$, see Draper and Pukelsheim (1999). Guan and Chao (1987) present the A-optimal design ξ_2^A in the second-degree model with $q \geq 4$, given by $\xi_2^A := \lambda_1\eta_1 + \lambda_2\eta_2$ with weights

$$\lambda_1 = \frac{\sqrt{4q-3}}{2(q-1)+\sqrt{4q-3}}\,, \qquad \lambda_2 = \frac{2(q-1)}{2(q-1)+\sqrt{4q-3}}\,.$$

We ignore the case $q \in \{2,3\}$. Following the same idea as in Section 3, our *model-robust A-criterion* is a convex combination of the A-criteria in the first- and second-degree models,

$$\widetilde{\Psi}_{\tilde{r}}^A(\xi) := \tilde{r}\frac{\operatorname{tr} M_1^{-1}(\xi)}{\operatorname{tr} M_1^{-1}(\xi_1^A)} + (1-\tilde{r})\frac{\operatorname{tr} M_2^{-1}(\xi)}{\operatorname{tr} M_2^{-1}(\xi_2^A)} \qquad \text{for } \tilde{r} \in [0,1].$$

From A-optimality results in the two individual models, we know $\operatorname{tr} M_1^{-1}(\xi_1^A) = q^2$ and $\operatorname{tr} M_2^{-1}(\xi_2^A) = 1/\lambda_1 + \binom{q}{2}\left[8/(q\lambda_1) + 16/(\binom{q}{2}\lambda_2)\right]$. These constants in $\widetilde{\Psi}_{\tilde{r}}^A(\cdot)$ standardize the A-criteria in the first- and second-degree model in order to make them comparable. Upon setting

$$r := r(\tilde{r}) := \frac{\tilde{r}/\operatorname{tr} M_1^{-1}(\xi_1^A)}{\tilde{r}/\operatorname{tr} M_1^{-1}(\xi_1^A) + (1-\tilde{r})/\operatorname{tr} M_2^{-1}(\xi_2^A)} \in [0,1], \tag{4.1}$$

we may rewrite $\widetilde{\Psi}_{\tilde{r}}^A(\xi)$ in the form

$$\widetilde{\Psi}_{\tilde{r}}^A(\xi) = \left(\frac{\tilde{r}}{\operatorname{tr} M_1^{-1}(\xi_1^A)} + \frac{1-\tilde{r}}{\operatorname{tr} M_2^{-1}(\xi_2^A)}\right)\left[r \operatorname{tr} M_1^{-1}(\xi) + (1-r)\operatorname{tr} M_2^{-1}(\xi)\right]$$

and thus eliminate the standardizing constants. Hence the design criterion

$$\Psi_r^A(\xi) := r \operatorname{tr} M_1^{-1}(\xi) + (1-r)\operatorname{tr} M_2^{-1}(\xi) \qquad \text{with } r \in [0,1] \tag{4.2}$$

is equivalent to $\widetilde{\Psi}_{\tilde{r}}^A$, and we use this simplified form. Note that the function $\tilde{r} \mapsto r(\tilde{r})$ from (4.1) is a bijection of $[0,1]$ into itself, which is why the transition from $\widetilde{\Psi}_{\tilde{r}}^A$ to $\Psi_r^A(\xi)$ is a mere re-parameterization of our family of model-robust A-criteria. For given $r \in [0,1]$, a design ξ^A with $M_2(\xi^A) \in \mathrm{PD}(m_2)$ is called *model-robust A-optimal* if and only if it satisfies

$$\Psi_r^A(\xi^A) = \min\{\Psi_r^A(\xi) \mid \xi \in \Xi \text{ with } M_2(\xi) \in \mathrm{PD}(m_2)\}.$$

The following lemma restricts the set of competing designs to the class $\mathcal{W}$ of weighted centroid designs. It parallels Lemma 3.1, and its proof is omitted.

Lemma 4.1. *The set $\mathcal{W}$ from Definition* 2.3 *is an essentially complete class of designs relative to the model-robust A-criterion Ψ_r^A, $r \in [0,1]$, defined in* (4.2). *Then a design ξ_r^A with $M_2(\xi_r^A) \in \mathrm{PD}(m_2)$ (for given $r \in [0,1]$) is model-robust A-optimal if and only if*

$$r \operatorname{tr} M_1^{-2}(\xi_r^A) M_1(\eta_j) + (1-r) \operatorname{tr} M_2^{-2}(\xi_r^A) M_2(\eta_j) \le r \operatorname{tr} M_1^{-1}(\xi_r^A) + (1-r) \operatorname{tr} M_2^{-1}(\xi_r^A)$$

for all $1 \le j \le q$.

We employ the same strategy as in Section 3. In our first step we consider designs that are convex combinations of the two A-optimal designs in the first- and second-degree model, and we establish a result analogous to that of Lemma 3.4.

Definition 4.2. We write $\Xi_A := \operatorname{conv}\{\xi_1^A, \xi_2^A\} \subseteq \mathcal{W}$ and $\xi_\alpha := \alpha \xi_1^A + (1-\alpha)\xi_2^A$ for all $\alpha \in [0,1]$.

Lemma 4.3. *For a given prior $r \in [0,1]$, there is a unique weight $\alpha_r^A \in [0,1]$ such that the design $\xi_{\alpha_r^A}$ is model-robust A-optimal among all designs in Ξ_A.*

Proof. For $\alpha \in [0,1]$, the information matrices of ξ_α are

$$M_1(\xi_\alpha) = \left[p_1 + \frac{1}{4}(q-1)p_2\right] I_q + \frac{1}{4} p_2 U_2 \,, \qquad M_2(\xi_\alpha) = \begin{pmatrix} M_1(\xi_\alpha) & \frac{1}{8} p_2 V_1' \\ \frac{1}{8} p_2 V_1 & \frac{1}{16} p_2 W_1 \end{pmatrix}$$

with

$$\begin{aligned} p_1 := p_1(\alpha) &:= \frac{\alpha}{q} + (1-\alpha)\lambda_1 = \frac{2(q-1)\alpha + \sqrt{4q-3}}{q[2(q-1)+\sqrt{4q-3}]}\,, \\ p_2 := p_2(\alpha) &:= (1-\alpha)\lambda_2 = \frac{4(1-\alpha)}{q[2(q-1)+\sqrt{4q-3}]}\,. \end{aligned}$$

The derivative of our model-robust A-criterion is

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\alpha} \Psi_r^A(\xi_\alpha) \\ &= r \operatorname{tr}\Big(M_1^{-2}(\xi_\alpha)[M_1(\xi_2^A) - M_1(\xi_1^A)]\Big) + (1-r)\operatorname{tr}\Big(M_2^{-2}(\xi_\alpha)[M_2(\xi_2^A) - M_2(\xi_1^A)]\Big). \end{aligned}$$

Lengthy algebra based on Klein's (2004a) multiplication table yields

$$M_1^{-2}(\xi_\alpha) = \bar{a}_1 U_1 + \bar{b}_1 U_2 \,, \qquad M_2^{-2}(\xi_\alpha) = \begin{pmatrix} \bar{a}_2 U_1 + \bar{b}_2 U_2 & \bar{c}_2 V_1' + \bar{d}_2 V_2' \\ \bar{c}_2 V_1 + \bar{d}_2 V_2 & \bar{e}_2 W_1 + \bar{f}_2 W_2 + \bar{g}_2 W_3 \end{pmatrix},$$

with

$$\begin{aligned}
\bar{a}_1 &= \frac{q^2\{[4p_1+(2q-3)p_2]^2+4(q-1)p_2^2\}}{[4p_1+(q-2)p_2]^2}\,, \\
\bar{b}_1 &= -\frac{q^2[8p_1+(3q-4)p_2]p_2}{[4p_1+(q-2)p_2]^2}\,, \\
\bar{a}_2 &= \frac{4q-3}{p_1^2}\,, \quad \bar{b}_2 = \frac{4}{p_1^2}\,, \qquad \bar{c}_2 = -\frac{2[16p_1+(4q+1)p_2]}{p_1^2p_2}\,, \quad \bar{d}_2 = -\frac{16}{p_1^2}\,, \\
\bar{e}_2 &= \frac{8[32p_1^2+32p_1p_2+(4q+1)p_2^2]}{p_1^2p_2^2}\,, \ \bar{f}_2 = \frac{4[32p_1+(4q+9)p_2]}{p_1^2p_2}\,, \qquad \bar{g}_2 = \frac{64}{p_1^2}\,.
\end{aligned}$$

Setting $(\mathrm{d}/\mathrm{d}\alpha)\Psi_r^A(\xi_\alpha)=0$ results in the equation $rt_1(\alpha)+(1-r)t_2(\alpha)=0$, where

$$\begin{aligned}
t_1(\alpha) &:= -\frac{q^2\,[\,2(q-1)+\sqrt{4q-3}\,]^2}{2[q\alpha+(q-2)+\sqrt{4q-3}]^2}, \\
t_2(\alpha) &:= q^2[2(q-1)+\sqrt{4q-3}]^2\left(\frac{1}{(1-\alpha)^2}-\frac{4q-3}{[\,2(q-1)\alpha+\sqrt{4q-3}\,]^2}\right),
\end{aligned}$$

and, solving for r, we obtain

$$r = h^{-1}(\alpha) := \frac{t_2(\alpha)}{t_2(\alpha)-t_1(\alpha)}\,. \tag{4.3}$$

It can be shown that $h^{-1}(0)=0$, $\lim_{\alpha\to 1}h^{-1}(\alpha)=1$, and $(\mathrm{d}/\mathrm{d}\alpha)h^{-1}(\alpha)>0$ for all $\alpha\in(0,1)$. Thus the function $h^{-1}:[0,1]\to[0,1]$ is bijective and, in particular, we have $r\in[0,1]$ in the above equation. The assertion follows by taking the inverse function h.

Figure 2 illustrates the behavior of the unique A-optimal weight α_r^A for various choices of q. Note that the above lemma does not provide an explicit description of the weight α_r^A since the equation $r=h^{-1}(\alpha)$ cannot be solved algebraically. Nevertheless we can now prove the design $\xi_{\alpha_r^A}$ to be model-robust A-optimal among *all* designs.

Theorem 4.4. *For a given* $r\in[0,1]$, *the design* $\xi_{\alpha_r^A}\in\Xi_A$ *with* $\alpha_r^A=h(r)\in[0,1]$ (*see Lemma* 4.3) *is model-robust A-optimal.*

Proof. For simplicity we start with an arbitrary weight $\alpha\in[0,1]$. Following Lemma 4.1 we investigate the quantities

$$N_{1,j}=\operatorname{tr} M_1^{-2}(\xi_\alpha)M_1(\eta_j)\,, \qquad N_{2,j}=\operatorname{tr} M_2^{-2}(\xi_\alpha)M_2(\eta_j), \qquad \text{for } 1\le j\le q.$$

The traces are computed using the multiplication table in Klein ((2004a), Lemma

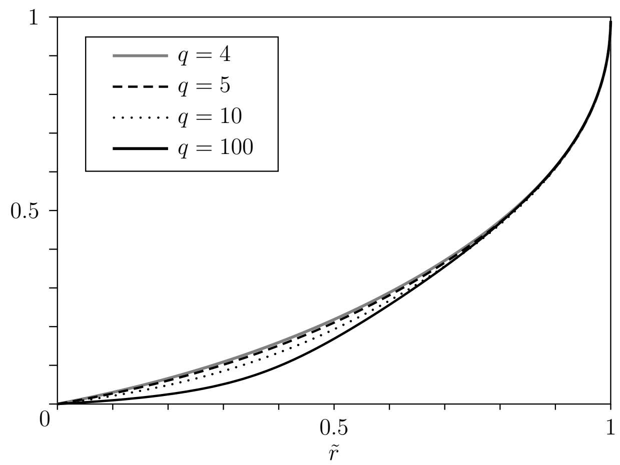

Figure 2. The A-optimal weight $\alpha^A_{r(\tilde{r})}$ as a function of $\tilde{r} \in [0,1]$. Lemma 4.3 yields a unique A-optimal weight α^A_r for every prior $r \in [0,1]$, that is, a function $r \mapsto \alpha^A_r$. Its graph is rather uninformative, which is why we use a parameterization in terms of $\tilde{r}$, defined by $r = r(\tilde{r})$ as in (4.1). Hence the function plotted above is $\tilde{r} \mapsto \alpha^A_{r(\tilde{r})}$. The weight $\alpha^A_{r(\tilde{r})}$ is found as the numerical solution of (4.3).

3.2) and the fact $\operatorname{tr} U_2 = \operatorname{tr} W_2 = \operatorname{tr} W_3 = 0$, resulting in

$$\begin{aligned} N_{1,j} &= \frac{\bar{a}_1 - \bar{b}_1}{j} + b_1\,, \\ N_{2,j} &= \frac{-\bar{e}_2/2 + 2\bar{f}_2 - 3\bar{g}_2/2}{j^3} + \frac{-2(\bar{c}_2 - \bar{d}_2) + \bar{e}_2/2 - 3\bar{f}_2 + 11\bar{g}_2/4}{j^2} \\ &\quad + \frac{\bar{a}_2 - \bar{b}_2 + 2(\bar{c}_2 - \bar{d}_2) - \bar{d}_2 + \bar{f}_2 - 3\bar{g}_2/2}{j} + \text{const.} \end{aligned}$$

with $\bar{a}_1, \bar{b}_1, \bar{a}_2, \ldots, \bar{g}_2$ as given in the proof of Lemma 4.3. The term $N_{1,j}$ is decreasing in j since $\bar{a}_1 - \bar{b}_1 > 0$. In order to prove $N_{2,j}$ decreasing in j, we find

$$\begin{aligned} N_{2,j} - N_{2,j+1} = \frac{1}{j^3(j+1)^3}\Big[&\Big(-\frac{1}{2}\,\bar{e}_2 + 2\bar{f}_2 - \frac{3}{2}\bar{g}_2\Big)(3j^2 + 3j + 1) \\ &+ \Big(-2(\bar{c}_2 - \bar{d}_2) + \frac{1}{2}\,\bar{e}_2 - 3\bar{f}_2 + \frac{11}{4}\,\bar{g}_2\Big)j(j+1)(2j+1) \\ &+ \Big(\bar{a}_2 - \bar{b}_2 + 2(\bar{c}_2 - \bar{d}_2) - \bar{d}_2 + \bar{f}_2 - \frac{3}{2}\,\bar{g}_2\Big)j^2(j+1)^2\Big] \end{aligned}$$

for $2 \le j \le q-1$. Due to $3j^2 + 3j + 1 \le j(j+1)(2j+1) \le j^2(j+1)^2$ for $j \ge 2$, we obtain

$$N_{2,j} - N_{2,\,j+1} \;\geq\; \frac{3j^2+3j+1}{j^3(j+1)^3}\Big(\bar{a}_2 - \bar{b}_2 - \bar{d}_2 - \frac{1}{4}\,\bar{g}_2\Big) = \frac{3j^2+3j+1}{j^3(j+1)^3}\cdot\frac{4q-7}{p_1^2} \;>\; 0.$$

Furthermore we have

$$r\,\mathrm{tr}\,M_1^{-2}(\xi_\alpha)M_1(\eta_1) + (1-r)\,\mathrm{tr}\,M_2^{-2}(\xi_\alpha)M_2(\eta_1) = r\bar{a}_1 + (1-r)\bar{a}_2 \tag{4.4}$$

with $\bar{a}_1, \bar{a}_2$ as given in the proof of Lemma 4.3, and

$$\begin{aligned} & r\,\mathrm{tr}\,M_1^{-1}(\xi_\alpha) + (1-r)\,\mathrm{tr}\,M_2^{-1}(\xi_\alpha) \\ &\quad = r\,\frac{q^2[4p_1+(2q-3)p_2]}{4p_1+(q-2)p_2} + (1-r)\frac{q[8(q-1)p_1+(4q-3)p_2]}{p_1p_2}\,. \end{aligned} \tag{4.5}$$

Now we consider the specific choice $\alpha := \alpha_r^A$, characterized as the unique solution of $r = t_2(\alpha)/[t_2(\alpha) - t_1(\alpha)]$ in (4.3). Using this identity, the right sides of (4.4) and (4.5) are seen to be equal, that is, we have equality in the optimality condition from Lemma 4.1 for $j = 1$. Finally, α_r^A satisfies $(\mathrm{d}/\mathrm{d}\alpha)\Psi_r^A(\xi_\alpha)|_{\alpha=\alpha_r^A} = 0$ by construction, which implies the terms $r\,\mathrm{tr}\,M_1^{-2}(\xi_{\alpha_r^A})M_1(\eta_j) + (1-r)\,\mathrm{tr}\,M_2^{-2}(\xi_{\alpha_r^A})M_2(\eta_j)$, $j = 1, 2$, coincide. This completes the proof.

5. Efficiencies of Model-Robust Optimal Designs

The model-robust D- and A-criteria used in Sections 3 and 4 contain a parameter $r \in [0,1]$ that represents a prior or weight in the tradeoff between the two candidate models $(\mathcal{S}^{q-1}, f_1)$ and $(\mathcal{S}^{q-1}, f_2)$. Since there is no natural choice for r, we are interested in finding a design $\xi_{\alpha_s^D}$ whose worst performance under all criteria Ψ_r^D, $r \in [0,1]$, is maximal. This maximin approach was suggested by Zen and Tsai (2002). Assume $r, s \in [0,1]$ fixed. The *relative* Ψ_r^D*-efficiency* of the model-robust D-optimal design $\xi_{\alpha_s^D}$ is defined as

$$D_r\text{-eff}(\xi_{\alpha_s^D}) := \frac{\Psi_r^D(\xi_{\alpha_s^D})}{\Psi_r^D(\xi_{\alpha_r^D})}\,, \tag{5.1}$$

where the optimal weight α_r^D is given in Lemma 3.4. The optimum value $\Psi_r^D(\xi_{\alpha_r^D})$ can be rewritten as

$$\begin{aligned} q(q+1)\,\Psi_r^D(\xi_{\alpha_r^D}) &= 2^{(1-q)[q\,(3-2r)+r]/[q(q+1)]}(q+1)^{r/q}(2+q+q\alpha_r^D)^{[(q-1)r]/q} \\ &\quad\times \Big((1-\alpha_r^D)^{(q-1)/2}[2-(1-q)\alpha_r^D]\Big)^{2(1-r)/(q+1)}\,. \end{aligned} \tag{5.2}$$

Substituting (5.2) into (5.1) we obtain $D_r\text{-eff}(\xi_{\alpha_s^D}) = g(r,s)/g(r,r)$ with

$$g(r,s) := (1-\alpha_s^D)^{(q-1)(1-r)/(q+1)}\,\big(2+(q-1)\alpha_s^D\big)^{(2-2r)/(q+1)}\,(2+q+q\alpha_s^D)^{(q-1)r/q}.$$

Now we are interested in finding a prior $s^* \in [0,1]$ such that

$$\min_{r\in[0,1]} D_r\text{-eff}(\xi_{\alpha^D_{s^*}}) \;=\; \max_{s\in[0,1]}\min_{r\in[0,1]} D_r\text{-eff}(\xi_{\alpha^D_s})\,. \tag{5.3}$$

The following lemma solves this maximin efficiency problem in two steps.

Theorem 5.1.

(i) *For fixed $s \in [0,1)$, the minimum $\min_{r\in[0,1]} D_r\text{-eff}(\xi_{\alpha^D_s})$ is attained at one of the endpoints $r=0$ and $r=1$. More precisely, we have*

$$\min_{r\in[0,1]} D_r\text{-eff}(\xi_{\alpha^D_s}) = \begin{cases} D_0\text{-eff}(\xi_{\alpha^D_s}) \textit{ for } s \ge s_1, \\ D_1\text{-eff}(\xi_{\alpha^D_s}) \textit{ for } s < s_1, \end{cases}$$

where $s_1 \in [0,1]$ is the unique solution of $p(s)=1$, with

$$p(s) := \frac{D_0\text{-eff}(\xi_{\alpha^D_s})}{D_1\text{-eff}(\xi_{\alpha^D_s})} = \frac{[2(q+1)]^{(q-1)/q}\Big((1-\alpha^D_s)^{(q-1)/2}[2+(q-1)\alpha^D_s]\Big)^{2/(q+1)}}{2^{2/(q+1)}(2+q+q\alpha^D_s)^{(q-1)/q}}\,.$$

(ii) *The choice $s^* = s_1$ satisfies the maximin efficiency property stated in* (5.3).

Proof. For convenience we treat $D_r\text{-eff}(\xi_{\alpha^D_s})$ as a function of α^D_r instead of r, utilizing the fact that $r \mapsto \alpha^D_r$ is a bijection. We have

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\alpha^D_r}\log D_r\text{-eff}(\xi_{\alpha^D_s}) \\ &= \frac{-2[(q^2-q-1)(\alpha^D_r)^2 + 2q\alpha^D_r + (q+2)]}{q\,(q+1)[2+(2q-1)\alpha^D_r + (\alpha^D_r)^2]^2} \\ &\quad \times \Big[(q-1)q\log\Big(\tfrac{1-\alpha^D_s}{1-\alpha^D_r}\Big) + 2q\log\Big(\tfrac{2+(q-1)\alpha^D_s}{2+(q-1)\alpha^D_r}\Big) - (q^2-1)\log\Big(\tfrac{2+q(1+\alpha^D_s)}{2+q(1+\alpha^D_r)}\Big)\Big]\,, \end{aligned}$$

which turns out to have a unique root between 0 and 1. Since $D_r\text{-eff}(\xi_{\alpha^D_s})$ attains a maximum at $\alpha^D_r = \alpha^D_s$ by definition, the two endpoints corresponding to $r=0$ and $r=1$ are the only local minima of $D_r\text{-eff}(\xi_{\alpha^D_s})$. In the case $s=s_1$, these two local minima are equal. Furthermore, we can show $(\mathrm{d}/\mathrm{d}\alpha^D_s)\log D_0\text{-eff}(\xi_{\alpha^D_s}) < 0$, which proves $D_0\text{-eff}(\xi_{\alpha^D_s})$ to be strictly decreasing in s. Similarly, $D_1\text{-eff}(\xi_{\alpha^D_s})$ is found to be strictly increasing. This establishes assertion (i). Part (ii) follows from the above monotonicity arguments.

Figure 3 displays the function $r \mapsto D_r\text{-eff}(\xi_{\alpha^D_s})$ for various choices of $s \in [0,1]$ and $q=5$. Table 1 lists numerical results for s^* and the corresponding maximin efficiencies of model-robust D-optimal designs, illustrating the results in Theorem 5.1.

Note that our concept of maximin efficiency only takes designs in the class Ξ_D into account. For any design ξ on $\mathcal{S}^{q-1}$, the remark following Lemma 3.1

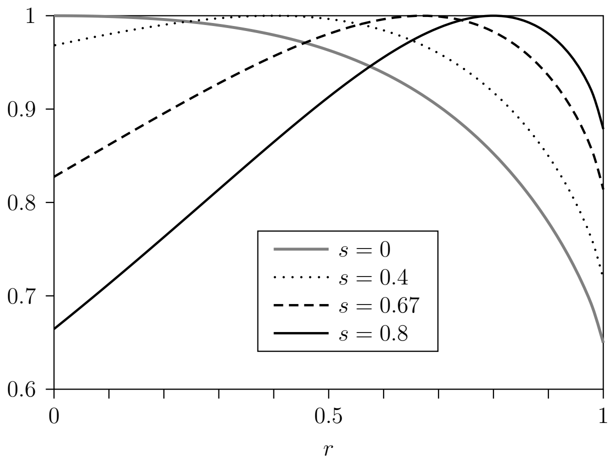

Figure 3. Efficiencies D_r-eff$(\xi_{\alpha_s^D})$ of model-robust D-optimal designs for $q = 5$.

Table 1. Maximin efficiencies $\min_{r\in[0,1]} D_r$-eff$(\xi_{\alpha_s^D})$ of model-robust D-optimal designs.

q	s^*	$\min_{r\in[0,1]} D_r$-eff$(\xi_{\alpha_{s^*}^D})$	D_0-eff$(\xi_{\alpha_{0.67}^D})$	D_1-eff$(\xi_{\alpha_{0.67}^D})$	D_1-eff$(\xi_{\alpha_0^D})$	D_0-eff$(\xi_{\alpha_{0.999}^D})$
2	0.679472	0.913557	0.915523	0.919615	0.816497	0.164968
3	0.679609	0.866132	0.869229	0.875693	0.731004	0.063190
4	0.679667	0.835551	0.839402	0.847453	0.681732	0.034628
5	0.679662	0.813938	0.818324	0.827503	0.649731	0.022876
10	0.679188	0.759294	0.764876	0.77658	0.579539	0.008436
100	0.672929	0.682612	0.685299	0.690824	0.508413	0.002444

The prior s^*, indicating the maximin-efficient model-robust D-optimal design from Theorem 5.1, converges to 0.67 as $q \to \infty$. We have D_0-eff$(\xi_{\alpha_s^D}) \approx D_1$-eff$(\xi_{\alpha_s^D})$ for $s = 0.67 \approx s^*$, where D_0-eff$(\xi_{\alpha_s^D})$ and D_1-eff$(\xi_{\alpha_s^D})$ are the candidates for an efficiency minimum. The two last columns show how poorly a design $\xi_{\alpha_s^D}$ may perform under variation of the criterion.

implies the existence of a weighted centroid design $\eta \in \mathcal{W}$ such that $\min_{r\in[0,1]} D_r$-eff$(\xi) \le \min_{r\in[0,1]} D_r$-eff$(\eta)$. Hence there is no need to consider minimum efficiencies of designs $\xi \notin \mathcal{W}$. To our knowledge, the existence of a weighted centroid design $\eta \in \mathcal{W}$ satisfying $\min_{r\in[0,1]} D_r$-eff$(\xi) > \min_{r\in[0,1]} D_r$-eff$(\xi_{\alpha_{s^*}^D})$ cannot be excluded in general.

Finally we consider efficiencies with respect to model-robust A-criteria. For $r, u \in [0,1]$, the *relative* Ψ_r^A*-efficiency* of the model-robust A-optimal design $\xi_{\alpha_u^A}$ is

$$A_r\text{-eff}(\xi_{\alpha_u^A}) := \frac{\Psi_r^A(\xi_{\alpha_r^A})}{\Psi_r^A(\xi_{\alpha_u^A})} = \frac{r\,\mathrm{tr}\,M_1^{-1}(\xi_{\alpha_r^A}) + (1-r)\mathrm{tr}\,M_2^{-1}(\xi_{\alpha_r^A})}{r\,\mathrm{tr}\,M_1^{-1}(\xi_{\alpha_u^A}) + (1-r)\mathrm{tr}\,M_2^{-1}(\xi_{\alpha_u^A})}\,,$$

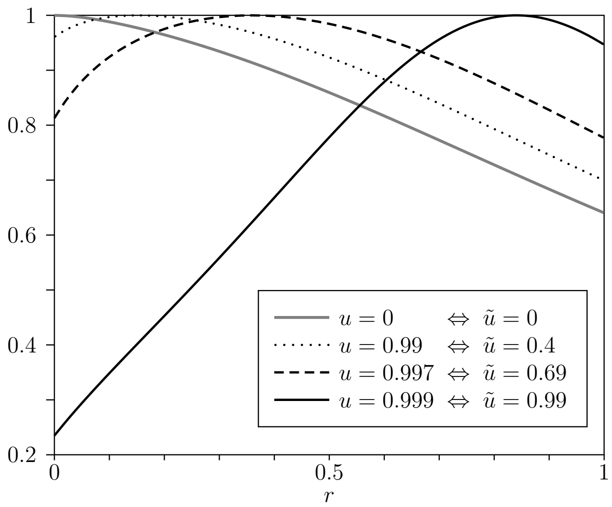

Figure 4. Efficiencies A_r-eff$(\xi_{\alpha_u^A})$ of model-robust A-optimal designs for $q = 5$.

Table 2. Maximin efficiencies $\min_{r\in[0,1]} A_r$-eff$(\xi_{\alpha_u^A})$ of model-robust A-optimal designs.

q	u^*	$\tilde{u}^*$	$\min_{r\in[0,1]} A_r$-eff$(\xi_{\alpha_{u^*}^A})$	A_1-eff$(\xi_{\alpha_{0.997}^A})$	A_0-eff$(\xi_{\alpha_{0.997}^A})$	A_1-eff$(\xi_{\alpha_0^A})$	A_0-eff$(\xi_{\alpha_{\approx 1}^A})$
4	0.995860	0.722769	0.797231	0.818912	0.737290	0.651388	$0.0^{3}2175$
5	0.997397	0.722747	0.786961	0.776796	0.812796	0.640388	$0.0^{3}2598$
10	0.999327	0.719108	0.760180	0.656392	0.968380	0.610099	$0.0^{3}4556$
100	0.999991	0.696118	0.704220	0.543873	0.999997	0.543621	0.0033872

The prior u^* indicates the maximin-efficient model-robust A-optimal design $\xi_{\alpha_{u^*}^A}$, in the sense of (5.4). The values of u^* are complemented with those of $\tilde{u}^*$, defined by $u^* = r(\tilde{u}^*)$ as in (4.1). While we have $u^* \to 1$ for $q \to \infty$, the weight $\tilde{u}^*$ clarifies the fact that the maximin efficient model-robust A-optimal design assigns considerable weight to ξ_2^A when $q \to \infty$.

where $\xi_{\alpha_r^A}$ are the designs from Definition 4.2. We are interested in $u^* \in [0,1]$ with

$$\min_{r\in[0,1]} A_r\text{-eff}(\xi_{\alpha_{u^*}^A}) = \max_{u\in[0,1]} \min_{r\in[0,1]} A_r\text{-eff}(\xi_{\alpha_u^A}) . \tag{5.4}$$

The arguments from Theorem 5.1 can be adapted to this notion of maximin efficiency. An outline of our numerical results on u^* for various choices of q is given in Table 2 and Figure 4.

6. Conclusion

In Section 2 we established a complete class result for mixed design criteria. The crucial argument is the compatibility of the Kiefer orderings of information matrices in the two candidate models, expressed by (1.2) and the identity

$ZH_\pi = R_\pi Z$, $\pi \in \mathfrak{S}_q$, in the proof of Lemma 2.1. Furthermore we derived model-robust D- and A-optimal designs (see Sections 3 and 4) that are convex combinations of the D- and A-optimal designs in the individual candidate models, respectively. This derivation was done in two steps, see the remark following Definition 3.3: First, we generated an optimality candidate in the class Ξ_D, then we showed this candidate to be optimal among *all* feasible designs. The particularly simple relation between optimal designs in the individual models and the model-robust optimal designs raises the question for a general condition ensuring such a smooth transition to model-robust designs. The case of model-robust A-optimality is more involved than its D-optimality analogue since the second-degree A-optimal design is not a uniform distribution on its support. Evaluating our equivalence theorems is tractable only due to the matrix algebra developed in Klein (2004a). Model-robust D- and A-optimal designs with maximin efficiency perform reasonably well with respect to the whole family of model-robust criteria considered here. At the same time, our results indicate that design performance may drop dramatically if the issue of efficiency under variation of the criterion is ignored.

Following our strategy, the model-robust optimality problem may be solved for other criteria than the D- and A-criteria investigated in this paper. The crucial requirement is that the selected optimality criteria are monotonic with respect to the Kiefer ordering, see Lemma 2.2. Nevertheless, alternative choices of optimality criteria will, in general, require extensive calculations and might involve intractable algebra. In principle our arguments also carry over to the setting where there is uncertainty between polynomial regression of degree two or three. However, no non-trivial essentially complete class of designs (as provided by Draper et al. (2000) for the second-degree model) is known for the third-degree model. As a result we do not have a completeness result analogous to Theorem 2.3 in the aforementioned setting, and model-robust D- and A-optimal designs are not available. Cook and Wong (1994) point out the equivalence of compound optimal designs (which is an alternative view of our model-robust design strategy) and constrained optimal designs under certain assumptions. Our treatment of model-robust design optimality is in line with this view and can therefore be used for further research on constrained optimality in mixture experiments.

Acknowledgement

The authors are grateful to the National Science Council (NSC) of Taiwan, R.O.C., and the Deutscher Akademischer Austausch Dienst who supported the research on this paper through their joint PPP program and through NSC grant 92-2118-M-110-003. The authors would also like to thank the anonymous referees for their insightful comments.

References

Atkinson, A. C. and Cox, D. R. (1974). Planning experiments for discriminating between models (with discussion). *J. Roy. Statist. Soc. Ser. B* **36**, 321-348.

Box, G. E. and Draper, N. R. (1959). A basis for the selection of a response surface design. *J. Amer. Statist. Assoc.* **54**, 622-654.

Chan, L.-Y. (2000). Optimal designs for experiments with mixtures: a survey. *Comm. Statist. Theory Methods* **29**, 2281-2312.

Cook, R. D. and Wong, W. K. (1994) . On the equivalence of constrained and compound optimal designs. *J. Amer. Statist. Assoc.* **89**, 687-692.

Cornell, J. A. (2002). *Experiments with Mixtures*. Third edition. Wiley, New York.

Dette, H. (1990). A generalization of D- and D_1-optimal designs in polynomial regression. *Ann. Statist.* **18**, 1784-1804.

Dette, H. (1991). A note on robust designs for polynomial regression. *J. Statist. Plann. Inference* **28**, 223-232.

Dette, H. (1993). On a mixture of the D- and D_1-optimality criterion in polynomial regression. *J. Statist. Plann. Inference* **35**, 233-249.

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.

Draper, N. R. and Pukelsheim, F. (1999). Kiefer ordering of simplex designs for first- and second-degree mixture models. *J. Statist. Plann. Inference* **79**, 325-348.

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

Galil, Z. and Kiefer, J. C. (1977). Comparison of simplex designs for quadratic mixture models. *Technometrics* **19**, 445-453.

Guan, Y. and Chao, X. (1987). On the A-optimal allocation of observations for the generalized simplex-centroid design (in Chinese). *J. Engineering Math.* **4**, 33-39.

Huang, M.-N. L. and Studden, W. J. (1988). Model robust extrapolation designs. *J. Statist. Plann. Inference* **18**, 1-24.

Kiefer, J. C. (1961). Optimum designs in regression problems, II. *Ann. Math. Statist.* **32**, 298-325.

Klein, T. (2004a). Invariant symmetric block matrices for the design of mixture experiments. *Linear Algebra and Its Applications* **388**, 261–278.

Klein, T. (2004b). Optimal designs for second-degree Kronecker model mixture experiments. *J. Statist. Plann. Inference* **123**, 117-131.

Läuter, E. (1974). Experimental design in a class of models. *Mathematische Operationsforschung und Statistik* **5**, 379-398.

Läuter, E. (1976). Optimal multipurpose designs for regression models. *Mathematische Operationsforschung und Statistik* **7**, 51-68.

Pukelsheim, F. (2006). *Optimal Design of Experiments*, Vol. 50 of *SIAM Classics in Applied Mathematics*. SIAM, Philadelphia, PA.

Pukelsheim, F. and Rosenberger, J. L. (1993). Experimental designs for model discrimination. *J. Amer. Statist. Assoc.* **88**, 642-649.

Scheffé, H. (1958). Experiments with mixtures. *J. Roy. Statist. Soc. Ser. B* **20**, 344-360.

Studden, W. J. (1982). Some robust-type D-optimal designs in polynomial regression. *J. Amer. Statist. Assoc.* **77**, 916-921.

Zen, M.-M. and Tsai, M.-H. (2002) . Some criterion-robust optimal designs for the dual problem of model discrimination and parameter estimation. *Sankhyā Ser. B* **64**, 322-338.

Department of Applied Mathematics, National Sun Yat-sen University, 70 Lien-hai Road, Kaohsiung 804, Taiwan, R.O.C.

E-mail: lomn@math.nsysu.edu.tw

Department of Applied Mathematics, National Sun Yat-sen University, 70 Lien-hai Road, Kaohsiung 804, Taiwan, R.O.C.

E-mail: hsuhl@math.nsysu.edu.tw

Department of Applied Mathematics, National Sun Yat-sen University, 70 Lien-hai Road, Kaohsiung 804, Taiwan, R.O.C.

E-mail: gchoucj@gmail.com

Zentrum Mathematik, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany.

E-mail: tpklein@ma.tum.de

(Received September 2006; accepted March 2008)