The Annals of Statistics
1995, Vol. 23, No. 1, 41–54

COMPLETE CLASS RESULTS FOR THE MOMENT MATRICES OF DESIGNS OVER PERMUTATION-INVARIANT SETS[1]

BY CHING-SHUI CHENG

University of California, Berkeley

In 1987 Cheng determined ϕ_p-optimal designs for linear regression (without intercept) over the n-dimensional unit cube $[0, 1]^n$ for $-\infty \le p \le 1$. These are uniform distributions on the vertices with a fixed number of entries equal to unity, and mixtures of neighboring such designs. In 1989 Pukelsheim showed that this class of designs is essentially complete and that the corresponding class of moment matrices is minimally complete, with respect to what he called Kiefer ordering. In this paper, these results are extended to general permutation-invariant design regions.

1. Introduction. Let X be a permutation-invariant and compact set in $\mathbb{R}_+^n \equiv \{\mathbf{x} = (x_1, \dots, x_n)^T\colon x_i \ge 0\}$, and let Ξ be the set of all the probability measures on the Borel subsets of X. Each $\xi \in \Xi$ is called an approximate design, or simply a design. For each $\xi \in \Xi$, define its moment matrix $\mathbf{M}(\xi)$ to be

$$\mathbf{M}(\xi) = \int_X \mathbf{x}\mathbf{x}^T \xi(d\mathbf{x}),$$

the $n \times n$ matrix whose (i, j)th entry is equal to $\int_X x_i x_j \xi(d\mathbf{x})$. Consider the linear regression model (without the constant term) on X:

$$E(y_{\mathbf{x}}) = \mathbf{x}^T\boldsymbol{\theta},$$

where $y_{\mathbf{x}}$ is an observation at $\mathbf{x} \in X$, and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)^T$ is the vector of unknown parameters. The observations are assumed to be uncorrelated with constant variance. Then $\mathbf{M}(\xi)$ is also the information matrix of ξ. This paper is concerned with the problem of optimally choosing ξ.

Pukelsheim (1989) introduced an ordering among such moment matrices. We say that a moment matrix $\mathbf{M}(\xi_1)$ is at least as informative as another moment matrix $\mathbf{M}(\xi_2)$ (or design ξ_1 is at least as informative as ξ_2), denoted by $\mathbf{M}(\xi_1) \gg_K \mathbf{M}(\xi_2)$, if there is an $n \times n$ matrix $\mathbf{B}$ (not necessarily a moment matrix) such that

(1.1) $$\mathbf{M}(\xi_1) - \mathbf{B} \text{ is nonnegative definite}$$

and

(1.2) $$\mathbf{B} \text{ belongs to the convex hull of } \{\mathbf{P}\mathbf{M}(\xi_2)\mathbf{P}^T\colon \mathbf{P} \in \mathrm{Perm}(n)\},$$

Received September 1993; revised May 1994.
[1]Research supported by NSF Grant DMS-91-00938.
AMS 1991 *subject classification*. Primary 62K05.
Key words and phrases. Essentially complete class, Kiefer ordering, minimally complete class, moment matrix, optimal design, Schur ordering.

where Perm(n) is the symmetric group consisting of all the $n \times n$ permutation matrices. Pukelsheim (1993) called $\gg_K$ the Kiefer ordering (relative to the symmetric group), apparently because the two steps (1.1) and (1.2) go back to Kiefer's (1975) fundamental result on universal optimality. A subset $\mathscr{D}$ of Ξ is called an essentially complete class with respect to this ordering if, for any $\xi \in \Xi$, there exists a $\xi' \in \mathscr{D}$ such that $\mathbf{M}(\xi') \gg_K \mathbf{M}(\xi)$. Furthermore, a moment matrix $\mathbf{M}(\xi_1)$ is more informative than another moment matrix $\mathbf{M}(\xi_2)$ if $\mathbf{M}(\xi_1) \gg_K \mathbf{M}(\xi_2)$, but $\mathbf{M}(\xi_1)$ is not of the form $\mathbf{P}\mathbf{M}(\xi_2)\mathbf{P}^T$, where $\mathbf{P} \in \mathrm{Perm}(n)$. A set $\mathscr{M}$ of moment matrices is minimally complete if, for any moment matrix $\mathbf{M}$ not in $\mathscr{M}$, there exists a moment matrix $\mathbf{M}^*$ in $\mathscr{M}$ such that $\mathbf{M}^*$ is more informative than $\mathbf{M}$ and there is no proper subset of $\mathscr{M}$ with the same property.

It follows from the preceding definitions that if ξ_1 is at least as informative as ξ_2, then $\phi(\mathbf{M}(\xi_1)) \geq \phi(\mathbf{M}(\xi_2))$ for all real-valued functions ϕ that are concave, permutation-invariant and Loewner-isotonic [here Loewner-isotonicity means that if $\mathbf{A} - \mathbf{B}$ is nonnegative definite, then $\phi(\mathbf{A}) \geq \phi(\mathbf{B})$]. We shall say that a design is ϕ-optimal if its moment matrix is positive definite and maximizes $\phi(\mathbf{M}(\xi))$ over $\xi \in \Xi$. Therefore, if $\mathscr{D}$ is essentially complete and there exists a ϕ-optimal design in Ξ, then there exists a design $\bar{\xi}$ in $\mathscr{D}$ such that $\bar{\xi}$ is also ϕ-optimal.

An important family of concave, permutation-invariant and Loewner-isotonic functions are the p-means ϕ_p. For any p such that $-\infty < p \leq 1$ and $p \neq 0$, let

$$\phi_p(\mathbf{M}) = \left[n^{-1}\,\mathrm{tr}\big(\mathbf{M}(\xi)^p\big)\right]^{1/p}.$$

Furthermore, define

$$\phi_0(\mathbf{M}) = \left[\det(\mathbf{M}(\xi))\right]^{1/n}$$

and

$$\phi_{-\infty}(\mathbf{M}) = \text{the smallest eigenvalue of } \mathbf{M}(\xi).$$

Then we have $\lim_{p \to -\infty} \phi_p = \phi_{-\infty}$ and $\lim_{p \to 0} \phi_p = \phi_0$. The ϕ_0-, ϕ_{-1}- and $\phi_{-\infty}$-optimal designs are the well known D-, A- and E-optimal designs, respectively.

One purpose of this paper is to determine ϕ_p-optimal designs for linear regression (without the constant term) over a general permutation-invariant set X in $\mathbb{R}^n_+$. This extends the work of Cheng (1987), who studied the n-dimensional unit cube $X = [0, 1]^n$, solving a conjecture of Harwit and Sloane (1976) in Hadamard transform optics. For $X = [0, 1]^n$, the ϕ-optimal designs have a very nice but curious structure: depending on the values of p, they are uniform distributions on the vertices with a fixed number of entries equal to 1, or are mixtures of neighboring such designs. Pukelsheim (1989) showed that this class of designs is essentially complete under the Kiefer ordering. Example 2 in Section 4 of the present article demonstrates that, for other sets X, the class of ϕ_p-optimal designs may not be essentially complete; so Pukelsheim's result cannot be extended to all permutation-invariant sets.

However, the extension holds for a somewhat different but closely related ordering which we call the Schur ordering. For any two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, we say that $\mathbf{x}$ is upper weakly majorized by $\mathbf{y}$, denoted by $\mathbf{x} \prec^W \mathbf{y}$, if $\sum_{i=1}^k x_{[i]} \geq \sum_{i=1}^k y_{[i]}$ for all $1 \leq k \leq n$, where $x_{[1]} \leq x_{[2]} \leq \cdots \leq x_{[n]}$ and $y_{[1]} \leq y_{[2]} \leq \cdots \leq y_{[n]}$ are the ordered components of $\mathbf{x}$ and $\mathbf{y}$, respectively. We say that a moment matrix $\mathbf{M}(\xi_1)$ is at least as informative as another moment matrix $\mathbf{M}(\xi_2)$ under the Schur ordering, denoted by $\mathbf{M}(\xi_1) \gg_S \mathbf{M}(\xi_2)$, if the vector of eigenvalues of $\mathbf{M}(\xi_1)$ is upper weakly majorized by that of $\mathbf{M}(\xi_2)$. If $\mathbf{M}(\xi_1)$ and $\mathbf{M}(\xi_2)$ have different eigenvalues and $\mathbf{M}(\xi_1) \gg_S \mathbf{M}(\xi_2)$, then $\mathbf{M}(\xi_1)$ is said to be more informative than $\mathbf{M}(\xi_2)$. It follows from this definition and from Theorem A.8 in Marshall and Olkin [(1979), Chapter 3] that if $\mathbf{M}(\xi_1) \gg_S \mathbf{M}(\xi_2)$, then $\phi(\mathbf{M}(\xi_1)) \geq \phi(\mathbf{M}(\xi_2))$ for all ϕ that are Schur-concave and componentwise increasing functions of the eigenvalues of the moment matrices. As in the case of Kiefer ordering, all the p-means ϕ_p, $-\infty \leq p \leq 1$, are covered. The concepts of essentially complete classes and minimally complete classes under the Schur ordering can be similarly defined.

In Section 2, essentially complete classes for Kiefer and Schur orderings for any permutation-invariant and compact set X in $\mathbb{R}_+^n$ are determined. In Section 3, ϕ_p-optimal designs are derived and are shown to constitute an essentially complete class for the Schur ordering. As a consequence, for any ϕ that is a Schur-concave and componentwise increasing function of the eigenvalues of the moment matrices, there exists a p, $-\infty \leq p \leq 1$, such that a certain ϕ_p-optimal design is also ϕ-optimal. As we mentioned earlier, this may not be true for the Kiefer ordering. Conditions on X under which the class of ϕ_p-optimal designs, $-\infty \leq p \leq 1$, is also essentially complete for the Kiefer ordering are given in Theorem 2.3. The cube $[0,1]^n$ is an example where Kiefer and Schur orderings give the same minimally complete class. While the derivation of ϕ_p-optimal designs in Cheng (1987) was based on the Kiefer–Wolfowitz equivalence theorem, the result on essentially complete classes provides an alternative solution. In light of the general results derived here, the curious structure of the optimal designs on the cube also becomes clearer. Detailed examples are given in Section 4.

For any $\mathbf{x} = (x_1, \ldots, x_n)^T \in X$ and any permutation π of $\{1, \ldots, n\}$, let $\pi(\mathbf{x})$ denote $(x_{\pi(1)}, \ldots, x_{\pi(n)})^T$. Throughout the rest of the paper, the orbit of $\mathbf{x}$, denoted $\langle \mathbf{x} \rangle$, is defined as the set $\{\pi(\mathbf{x})$: π is a permutation of $1, \ldots, n\}$. The uniform measure on $\langle \mathbf{x} \rangle$, that is, the measure which assigns equal weights to each $\pi(\mathbf{x})$, will be denoted by $\xi_{\langle \mathbf{x} \rangle}$. Finally, $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ $(= \sum_{i=1}^n x_i$ for $\mathbf{x} \in \mathbb{R}_+^n)$ and $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$ are the usual L_1- and L_2-norms.

2. Essentially complete classes. For any $\xi \in \Xi$, let $\bar{\xi}$ be the symmetrized version of ξ, that is,

$$\bar{\xi}(B) = \sum_{\pi} \frac{1}{n!} \xi(\pi(B)),$$

for all Borel subsets B, where the sum is over the $n!$ permutations of $\{1,\dots,n\}$. Then $\bar{\xi}$ is permutation-invariant and its moment matrix is also permutation-invariant. It is easy to see that $\bar{\xi}$ is at least as informative as ξ under both Kiefer and Schur orderings. Therefore the set of all permutation-invariant moment matrices (or moment matrices of permutation-invariant designs) is essentially complete for both orderings. Although this provides a substantial reduction of the problem, the set of permutation-invariant designs is still too large. We shall look for a further reduction, with the eventual goal to find a set of designs whose moment matrices are minimally complete. From now on, without loss of generality, all the designs ξ are assumed to be permutation-invariant.

If ξ is permutation-invariant, then $\mathbf{M}(\xi)$ is of the form $\alpha(\xi)\mathbf{I} + \beta(\xi)\mathbf{J}$, where $\mathbf{I}$ is the $n \times n$ identity matrix and $\mathbf{J}$ is the $n \times n$ matrix of 1's. Therefore ξ has at most two distinct eigenvalues: $\mu(\xi) = \alpha(\xi) + n\beta(\xi)$ with multiplicity 1, and $\nu(\xi) = \alpha(\xi)$ with multiplicity $n - 1$. Since $X \subset \mathbb{R}^n_+$, it is clear that $\beta(\xi) \geq 0$, and

$$\mu(\xi) \geq \nu(\xi). \tag{2.1}$$

For the Kiefer ordering, a permutation-invariant design ξ_1 is at least as informative as another permutation-invariant design ξ_2 if and only if

$$\mu(\xi_1) \geq \mu(\xi_2) \quad \text{and} \quad \nu(\xi_1) \geq \nu(\xi_2); \tag{2.2}$$

while for the Schur ordering the condition is

$$\begin{gathered} \nu(\xi_1) \geq \nu(\xi_2) \quad \text{and} \\ (n-1)\nu(\xi_1) + \mu(\xi_1) \geq (n-1)\nu(\xi_2) + \mu(\xi_2). \end{gathered} \tag{2.3}$$

In both cases, the comparison of permutation-invariant designs can be based on their two eigenvalues. Consequently, it is enough to consider the following subset of $\mathbb{R}^2$:

$$\mathscr{R} = \{(\mu(\xi), \nu(\xi)) \colon \xi \text{ is a permutation-invariant design in } \Xi\}.$$

The comparison of information matrices is thereby reduced to a two-dimensional problem.

The following can be said about $\mathscr{R}$: (a) $\mathscr{R}$ is compact and convex. (b) For any a which is equal to $\mu(\xi)$ for a certain ξ, let $g(a) = \max_{\xi:\, \mu(\xi)=a} \nu(\xi)$. If ξ^* is such that $\mu(\xi^*) = a$ and $\nu(\xi^*) = g(a)$, then ξ^* is at least as informative as ξ under both orderings.

In (a), the convexity follows from the fact that

$$\mu\left(\sum_{i=1}^{k} \alpha_i \xi_i\right) = \sum_{i=1}^{k} \alpha_i \mu(\xi_i) \quad \text{and} \quad \nu\left(\sum_{i=1}^{k} \alpha_i \xi_i\right) = \sum_{i=1}^{k} \alpha_i \nu(\xi_i)$$

for any convex combination $\sum_{i=1}^{k} \alpha_i \xi_i$, and (b) is trivial. By (b), the moment matrices whose eigenvalues correspond to the upper boundary of $\mathscr{R}$ form an essentially complete class, that is, $\{\xi \colon \xi \text{ is permutation-invariant and } \nu(\xi) = g(\mu(\xi))\}$ is an essentially complete class for both orderings. This further reduces the comparison of information matrices to a one-dimensional problem and is also where the two orderings may start to give different results. We now consider Kiefer and Schur orderings separately.

2.1. *Kiefer ordering*. Since $\mathscr{R}$ is convex, the function g that defines its upper boundary is concave throughout the interval $[\min_\xi \mu(\xi), \max_\xi \mu(\xi)]$. Let A be the set of all the a's at which the maximum of $g(a)$ over $[\min_\xi \mu(\xi), \max_\xi \mu(\xi)]$ is attained, and let a_L be the largest number in A. Furthermore, let $a_R = \max_\xi \mu(\xi)$. Then

$$\text{(2.4)} \qquad \begin{array}{l} g \text{ is increasing on } \left[\min_\xi \mu(\xi), a_L\right] \\ \text{and strictly decreasing on } [a_L, a_R]. \end{array}$$

An immediate consequence of (2.1) and (2.4) is

$$\text{(2.5)} \qquad a > g(a) \quad \text{for all } a \in (a_L, a_R], \quad \text{and} \quad a_L \geq g(a_L).$$

Let

$$\mathscr{R}_K = \{(a, b): a_L \leq a \leq a_R, b = g(a)\}.$$

Then the following lemma shows that $\mathscr{R}_K$ is the part of the upper boundary of $\mathscr{R}$ that corresponds to the eigenvalues of the moment matrices in a minimally complete class.

LEMMA 2.1. *For linear regression without the constant term over any permutation-invariant and compact set X in $\mathbb{R}^n_+$, an essentially complete class of designs for the Kiefer ordering is*

$$\mathscr{C}_K = \{\xi: \xi \text{ is permutation-invariant and } (\mu(\xi), \nu(\xi)) \in \mathscr{R}_K\},$$

and the moment matrices of the designs in $\mathscr{C}_K$ constitute a minimally complete class.

PROOF. Let ξ_L be a design such that $\mu(\xi_L) = a_L$ and $\nu(\xi_L) = g(a_L)$. By (2.4), for any ξ,

$$\text{(2.6)} \qquad \text{if } \mu(\xi) < a_L, \text{ then } \nu(\xi) \leq g(a_L).$$

Therefore ξ_L is more informative than ξ under the Kiefer ordering. It follows that all the designs with $\mu(\xi) < a_L$ can be eliminated. Hence $\mathscr{C}_K$ is an essentially complete class under the Kiefer ordering. Furthermore, since g is strictly decreasing on $[a_L, a_R]$, for any two designs ξ_1 and ξ_2 in $\mathscr{C}_K$, if $\mu(\xi_1) < \mu(\xi_2)$, then $\nu(\xi_1) > \nu(\xi_2)$. In other words, neither design is more informative than the other. This shows that the moment matrices of the designs in $\mathscr{C}_K$ constitute a minimally complete class under the Kiefer ordering. □

Since $\nu(\xi)$ is the smallest eigenvalue of $\mathbf{M}(\xi)$, by the definitions of a_L and ξ_L, we see that ξ_L is an E-optimal design over Ξ. [If the set A contains more than one number, then any ξ with $\mu(\xi) \in A$ and $\nu(\xi) = g(\mu(\xi))$ is also E-optimal, but ξ_L is at least as informative as any such ξ.]

For convenience, denote the point $(a, g(a))$ by $\mathbf{g}(a)$. Then $\mathscr{R}_K$, consisting of the points $\mathbf{g}(a)$ with $a_L \le a \le a_R$, is a curve whose leftmost point $\mathbf{g}(a_L)$ corresponds to an E-optimal design.

It is clear that each point in $\mathscr{R}_K$ is a convex combination of *at most two* extreme points of $\mathscr{R}$. Although extreme points are usually defined for convex sets only, one can extend the definition to nonconvex sets; so $\mathbf{x}$ is an extreme point of a set S if it cannot be expressed as $\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$, where $0 < \alpha < 1$ and $\mathbf{x}_1, \mathbf{x}_2 \in S$. With this in mind, we can say that each point in $\mathscr{R}_K$ is a convex combination of at most two extreme points of $\mathscr{R}_K$. More precisely, if $\mathbf{g}(a) \in \mathscr{R}_K$ is not an extreme point, then it lies on the line segment formed by the two neighboring extreme points.

We shall give a further reduction of $\mathscr{C}_K$ to a much smaller subclass which can be described in terms of the geometry of X. There are many permutation-invariant designs corresponding to the same point of $\mathscr{R}_K$, and it is enough to keep only one of them. Let $\tilde{X} = \{\mathbf{x} \in X\colon \mathbf{x}$ is an extreme point of the convex hull of $X \cup (-X)\}$. It follows from Theorem 8.5 of Pukelsheim [(1993), page 191] and the discussion in the preceding paragraph that each design in $\mathscr{C}_K$ has the same moment matrix as $\xi_{\langle\mathbf{x}\rangle}$ for some $\mathbf{x} \in \tilde{X}$ (which corresponds to an extreme point of $\mathscr{R}_K$), or a convex combination of two $\xi_{\langle\mathbf{x}\rangle}$'s, say, $\xi_{\langle\mathbf{x}_1\rangle}$ and $\xi_{\langle\mathbf{x}_2\rangle}$, where $\mathbf{x}_1, \mathbf{x}_2 \in \tilde{X}$. Our final task is to characterize the points $\mathbf{x} \in \tilde{X}$ such that the moment matrices of $\xi_{\langle\mathbf{x}\rangle}$ or convex combinations of two such designs belong to a minimally complete class.

A simple calculation shows that for each $\mathbf{x} \in X$, we have $\mathbf{M}(\xi_{\langle\mathbf{x}\rangle}) = \alpha_{\mathbf{x}}\mathbf{I} + \beta_{\mathbf{x}}\mathbf{J}$, with $\alpha_{\mathbf{x}} = (n\|\mathbf{x}\|_2^2 - \|\mathbf{x}\|_1^2)/[n(n-1)]$ and $\beta_{\mathbf{x}} = (\|\mathbf{x}\|_1^2 - \|\mathbf{x}\|_2^2)/[n(n-1)]$. So

$$\mu(\xi_{\langle\mathbf{x}\rangle}) = \alpha_{\mathbf{x}} + n\beta_{\mathbf{x}} = n^{-1}\|\mathbf{x}\|_1^2 \tag{2.7}$$

and

$$\nu(\xi_{\langle\mathbf{x}\rangle}) = \alpha_{\mathbf{x}} = \left(\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2\right)/(n-1). \tag{2.8}$$

Notice that $n^{-1}\|\mathbf{x}\|_1^2$ and $\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2$ are the squared lengths of the orthogonal projections of $\mathbf{x}$ onto the equiangular line and its orthogonal complement, respectively. Let $\mathbf{x}_L$ maximize $\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2$ over $\mathbf{x} \in \tilde{X}$, and maximize $\|\mathbf{x}\|_1^2$ among those which maximize $\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2$; furthermore, let $\mathbf{x}_R$ maximize $\|\mathbf{x}\|_1^2$ over $\mathbf{x} \in \tilde{X}$, and maximize $\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2$ (or, equivalently, maximize $\|\mathbf{x}\|_2$) among those which maximize $\|\mathbf{x}\|_1^2$. Then by (2.7), (2.8) and the definitions of a_L and a_R, it is clear that $\xi_{\langle\mathbf{x}_L\rangle}$ and $\xi_{\langle\mathbf{x}_R\rangle}$ correspond to the leftmost and rightmost points of $\mathscr{R}_K$, respectively, and $a_L = n^{-1}\|\mathbf{x}_L\|_1^2$ and $a_R = n^{-1}\|\mathbf{x}_R\|_1^2$.

For each a with $n^{-1}\|\mathbf{x}_L\|_1^2 < a < n^{-1}\|\mathbf{x}_R\|_1^2$, we would like to find a design corresponding to $\mathbf{g}(a)$ in the form of $\xi_{\langle\mathbf{x}\rangle}$, where $\mathbf{x} \in \tilde{X}$, or a convex combination of two such designs. Let C_a be the set

$$\left\{\mathbf{y} \in \tilde{X}\colon \mathbf{y} \text{ maximizes } \|\mathbf{x}\|_2 \text{ over all the points of } \tilde{X} \text{ with } n^{-1}\|\mathbf{x}\|_1^2 = a\right\}.$$

Pick an arbitrary point from each C_a if it is nonempty. Let E be the resulting set, and let E_K be

$$\left\{\mathbf{x} \in E\colon \left(n^{-1}\|\mathbf{x}\|_1^2, \left(\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2\right)/(n-1)\right) \text{ is an extreme point of } \mathscr{R}_K\right\}.$$

Then we have the following theorem.

THEOREM 2.1. *An essentially complete class for the Kiefer ordering is the set* $\mathscr{M}_K$ *consisting of the designs* ξ *such that* $\xi = \xi_{\langle \mathbf{x} \rangle}$, *for some* $\mathbf{x} \in E_K$, *or is a convex combination of* $\xi_{\langle \mathbf{x}_1 \rangle}$ *and* $\xi_{\langle \mathbf{x}_2 \rangle}$, *where* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are two neighboring points in* E_K *in the sense that there is no other point* $\mathbf{x} \in E_K$ *such that* $\|\mathbf{x}_1\|_1 < \|\mathbf{x}\|_1 < \|\mathbf{x}_2\|_1$ *or* $\|\mathbf{x}_2\|_1 < \|\mathbf{x}\|_1 < \|\mathbf{x}_1\|_1$.

PROOF. If $\exists\, \mathbf{y} \in \tilde{X}$ such that $\xi_{\langle \mathbf{y} \rangle}$ corresponds to $\mathbf{g}(a)$, then we must have $n^{-1}\|\mathbf{y}\|_1^2 = a$, and $\mathbf{y}$ must maximize $\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2$ (or, equivalently, maximize $\|\mathbf{x}\|_2$) over the set $\{\mathbf{x} \in \tilde{X}\colon n^{-1}\|\mathbf{x}\|_1^2 = a\}$. In other words, if $\xi_{\langle \mathbf{y} \rangle}$ corresponds to a point in $\mathscr{R}_K$, then $\exists\, \mathbf{x} \in E$ such that $\xi_{\langle \mathbf{y} \rangle}$ and $\xi_{\langle \mathbf{x} \rangle}$ have the same moment matrices. A $\xi_{\langle \mathbf{x} \rangle}$ with $\mathbf{x} \in E$ may not correspond to a point in $\mathscr{R}_K$, but it is clear that $\mathscr{R}_K$ is the upper boundary of the convex hull of the set $\{(n^{-1}\|\mathbf{x}\|_1^2, (\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2)/(n-1))\colon \mathbf{x} \in E\}$. If we remove all the points $\mathbf{x}$ in E such that $(n^{-1}\|\mathbf{x}\|_1^2, (\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2)/(n-1))$ is not an extreme point of $\mathscr{R}_K$, then the resulting set E_K has the property that each $\xi_{\langle \mathbf{x} \rangle}$ with $\mathbf{x} \in E_K$ corresponds to an extreme point of $\mathscr{R}_K$. Each of the remaining points of $\mathscr{R}_K$ lies on the line segment connecting two neighboring extreme points and corresponds to a convex combination of two neighboring $\xi_{\langle \mathbf{x} \rangle}$'s. □

The class $\mathscr{M}_K$ is obtained by picking one design out of those corresponding to the same point in $\mathscr{R}_K$. Thus there is a one-to-one correspondence between $\mathscr{M}_K$ and $\mathscr{R}_K$ through the mapping

$$\varepsilon \xi_{\langle \mathbf{x}_1 \rangle} + (1-\varepsilon)\xi_{\langle \mathbf{x}_2 \rangle} \to \mathbf{g}(a), \tag{2.9}$$

where $a = n^{-1}[\varepsilon\|\mathbf{x}_1\|_1^2 + (1-\varepsilon)\|\mathbf{x}_2\|_1^2]$, and $\mathbf{x}_1$ and $\mathbf{x}_2$ are two neighboring points in E_K.

2.2. *Schur ordering*. We shall show that the minimally complete class for the Schur ordering is a subclass of that for the Kiefer ordering.

LEMMA 2.2. *Let* a_R^S *be the smallest value of* a *that maximizes* $a + (n-1)g(a)$ *over* $a \in [a_L, a_R]$, *and let* $\mathscr{R}_S$ *be the subset* $\{\mathbf{g}(a)\colon a \in [a_L, a_R^S]\}$ *of* $\mathscr{R}_K$. *Then the moment matrices whose eigenvalues correspond to the points in* $\mathscr{R}_S$ *constitute a minimally complete class for the Schur ordering.*

PROOF. We have already seen that the moment matrices whose eigenvalues correspond to the upper boundary of $\mathscr{R}$ form an essentially complete class under both orderings. For any ξ with $\mu(\xi) < a_L = \mu(\xi_{\langle \mathbf{x}_L \rangle})$, by (2.4), we also have $\nu(\xi) \le g(a_L) = \nu(\xi_{\langle \mathbf{x}_L \rangle})$. It follows from (2.3) that $\xi_{\langle \mathbf{x}_L \rangle}$ is more informative than ξ under the Schur ordering. Therefore, all the designs with $\mu(\xi) < a_L$ can be eliminated.

Suppose ξ_R^S is a design with $\mu(\xi_R^S) = a_R^S$ and $\nu(\xi_R^S) = g(a_R^S)$. Then for any design ξ with $\mu(\xi) > a_R^S$, by (2.4), we must have $\nu(\xi) < \nu(\xi_R^S)$. Then since $(n-1)\nu(\xi_R^S) + \mu(\xi_R^S) \geq (n-1)\nu(\xi) + \mu(\xi)$, by (2.3), ξ_R^S is more informative than ξ. This eliminates all the points $\mathbf{g}(a)$ with $a > a_R^S$.

To finish the proof, we need to show that if ξ_1 and ξ_2 are two designs such that $a_L \leq \mu(\xi_1) < \mu(\xi_2) \leq a_R^S$, and $\nu(\xi_i) = g(\mu(\xi_i))$, $i = 1, 2$, then neither is more informative than the other. Since the function $a + (n-1)g(a)$ is concave and achieves its first maximum at a_R^S, it is strictly increasing on $[a_L, a_R^S]$. Therefore $\mu(\xi_1) + (n-1)\nu(\xi_1) < \mu(\xi_2) + (n-1)\nu(\xi_2)$. On the other hand, by (2.4), $\nu(\xi_1) > \nu(\xi_2)$. Hence neither of ξ_1 and ξ_2 is more informative than the other. □

Lemma 2.2 shows that $\mathscr{R}_S$ is a part of the curve $\mathscr{R}_K$, starting from $\mathbf{g}(a_L)$ and ending at $\mathbf{g}(a_R^S)$. For any design ξ corresponding to $\mathbf{g}(a)$, $a + (n-1)g(a) = \operatorname{tr}\mathbf{M}(\xi) \propto \phi_1(\mathbf{M}(\xi))$. Therefore the rightmost point of $\mathscr{R}_S$ corresponds to a ϕ_1-optimal design. From the mapping in (2.9), an essentially complete class for the Schur ordering can be obtained as a subclass of the essentially complete class $\mathscr{M}_K$ for the Kiefer ordering given in Theorem 2.1.

THEOREM 2.2. *Suppose* $\mathbf{x}_R^S$ *maximizes* $\|\mathbf{x}\|_2$ *over* $\mathbf{x} \in \tilde{X}$ *and also minimizes* $\|\mathbf{x}\|_1$ *among those which maximize* $\|\mathbf{x}\|_2$. *Let* E_S *be the following subset of* E_K:

$$E_S \equiv \{\mathbf{x} : \mathbf{x} \in E_K \textit{ and } \|\mathbf{x}\|_1 \leq \|\mathbf{x}_R^S\|_1\}.$$

Then an essentially complete class for the Schur ordering is

$$\begin{aligned}\mathscr{M}_S \equiv \{\xi : \xi = \xi_{\langle \mathbf{x} \rangle} & \textit{ for some } \mathbf{x} \in E_S, \\ & \textit{or is a convex combination of } \xi_{\langle \mathbf{x}_1 \rangle} \textit{ and } \xi_{\langle \mathbf{x}_2 \rangle}, \\ & \textit{where } \mathbf{x}_1 \textit{ and } \mathbf{x}_2 \textit{ are two neighboring points in } E_S\}.\end{aligned}$$

PROOF. By (2.7) and (2.8), $(n-1)\nu(\xi_{\langle \mathbf{x} \rangle}) + \mu(\xi_{\langle \mathbf{x} \rangle}) = \|\mathbf{x}\|_2^2$. So by Lemma 2.2, a_R^S can be obtained by maximizing $\|\mathbf{x}\|_2$ over $\mathbf{x} \in \tilde{X}$, and then minimizing $\|\mathbf{x}\|_1$ among those which maximize $\|\mathbf{x}\|_2$. Then $a_R^S = n^{-1}\|\mathbf{x}_R^S\|_1^2$, and $\xi_{\langle \mathbf{x}_R^S \rangle}$ corresponds to $\mathbf{g}(a_R^S)$, the rightmost point of $\mathscr{R}_S$. □

The following result is concerned with when minimally complete classes for the two orderings are the same.

THEOREM 2.3. *The minimally complete classes of moment matrices for Kiefer and Schur orderings are the same if and only if all the points that maximize* $\|\mathbf{x}\|_2$ *over* $\tilde{X}$ *also maximize* $\|\mathbf{x}\|_1$ (*therefore all the points maximizing* $\|\mathbf{x}\|_2$ *over* $\tilde{X}$ *have the same* L_1-*norm*).

PROOF. Comparing $\mathscr{R}_K$ with $\mathscr{R}_S$, we see that the two orderings have the same minimally complete classes of moment matrices if and only if $a_R = a_R^S$, or, equivalently,

$$\|\mathbf{x}_R^S\|_1 = \|\mathbf{x}_R\|_1. \tag{2.10}$$

By definition, $\mathbf{x}_R^S$ maximizes $\|\mathbf{x}\|_2$ over $\mathbf{x} \in \tilde{X}$, and minimizes $\|\mathbf{x}\|_1$ among those which maximize $\|\mathbf{x}\|_2$. On the other hand, $\mathbf{x}_R$ maximizes $\|\mathbf{x}\|_1$ over $\mathbf{x} \in \tilde{X}$, and maximizes $\|\mathbf{x}\|_2$ among those which maximize $\|\mathbf{x}\|_1$. It is easy to see that (2.10) is equivalent to all the points that maximize $\|\mathbf{x}\|_2$ over $\tilde{X}$ also maximizing $\|\mathbf{x}\|_1$. □

Detailed examples illustrating the results in this section can be found in Section 4.

3. ϕ_p-optimal designs. Since the ϕ_p-criteria are Schur concave, ϕ_p-optimal designs can be found among the designs in the essentially complete class $\mathscr{M}_S$. In other words, for any $-\infty \le p \le 1$, there exists a ϕ_p-optimal design of the form $\xi_{\langle \mathbf{x} \rangle}$ for some $\mathbf{x} \in E_S$, or a convex combination of $\xi_{\langle \mathbf{x}_1 \rangle}$ and $\xi_{\langle \mathbf{x}_2 \rangle}$, where $\mathbf{x}_1$ and $\mathbf{x}_2$ are two neighboring points in E_S. We shall show that the converse is also true.

THEOREM 3.1. *Each design in $\mathscr{M}_S$ is ϕ_p-optimal for at least one $p \in [-\infty, 1]$.*

PROOF. For $-\infty < p \le 1$, $p \ne 0$, define a function f_p on $[a_L, a_R^S]$ by

$$f_p(a) = \left(n^{-1}\{a^p + (n-1)[g(a)]^p\}\right)^{1/p};$$

for $p = 0$, let $f_0(a) = \{a \cdot [g(a)]^{n-1}\}^{1/n}$; and for $p = -\infty$, let $f_{-\infty}(a) = g(a)$. To determine a ϕ_p-optimal design, we can maximize $f_p(a)$ over $a \in [a_L, a_R^S]$. If the maximum is attained at a, then the design in $\mathscr{M}_S$ corresponding to $\mathbf{g}(a)$ is ϕ_p-optimal over Ξ. Due to the one-to-one correspondence between $\mathscr{M}_S$ and $\mathscr{R}_S$, for simplicity, in the rest of the proof, we shall only deal with $\mathscr{R}_S$. If the maximum of f_p is attained at a, then we say that $\mathbf{g}(a)$ is ϕ_p-optimal. It is enough to show that each $\mathbf{g}(a)$, $a \in [a_L, a_R^S]$, is ϕ_p-optimal for at least one $p \in [-\infty, 1]$.

Since g is concave on $[a_L, a_R^S]$, for all $-\infty < p < 1$, f_p is strictly concave on $[a_L, a_R^S]$. Therefore the maximum of $f_p(a)$ over $[a_L, a_R^S]$ is achieved at a unique a. By the definitions of a_L and a_R^S, this is also true for $p = -\infty$ and 1. Let

$$F(p) \equiv \max_{a \in [a_L, a_R^S]} f_p(a).$$

Using the strict concavity of f_p for all $-\infty < p < 1$ and the fact that, for fixed a, $f_p(a)$ is a continuous function of $p \in [-\infty, 1]$, it can easily be seen that $F(p)$ is a continuous function of $p \in [-\infty, 1]$. We have already known that $\mathbf{g}(a_L)$ is $\phi_{-\infty}$-optimal and $\mathbf{g}(a_R^S)$ is ϕ_1-optimal. The continuity of $F(p)$ over $p \in [-\infty, 1]$ now implies that each $\mathbf{g}(a)$, $a \in [a_L, a_R^S]$, is ϕ_p-optimal for at least one $p \in [-\infty, 1]$. □

A consequence of Theorem 3.1 is that the ϕ_p-optimal designs constitute an essentially complete class for the Schur ordering. When the conditions in Theorem 2.3 are satisfied, they are also an essentially complete class for the Kiefer ordering.

When $a_L = a_R^S$, $\mathscr{R}_S$ consists of one single point, and $\mathbf{g}(a_L)$ is ϕ_p-optimal for all $p \in [-\infty, 1]$. Therefore, to derive ϕ_p-optimal designs, without loss of generality, we may assume that $a_L < a_R^S$. For any $-\infty < p < 1$, since $f_p(a)$ is a concave function of a, the maximum of f_p over $[a_L, a_R^S]$ occurs at $a \in (a_L, a_R^S)$ if and only if

$$f_p'(a_+) \le 0 \tag{3.1}$$

and

$$f_p'(a_-) \ge 0, \tag{3.2}$$

where $f_p'(a_+)$ and $f_p'(a_-)$ are the right- and left-derivatives of f_p at a, respectively. Similarly, the maximum occurs at a_L (respectively, a_R^S) if and only if $f_p'(a_{L+}) \le 0$ [respectively, $f_p'(a_{R-}^S) \ge 0$].

By direct calculation, (3.1) and (3.2) are equivalent to

$$[a/g(a)]^{p-1} \le -(n-1)g'(a_+) \tag{3.3}$$

and

$$[a/g(a)]^{p-1} \ge -(n-1)g'(a_-). \tag{3.4}$$

We have the following theorem.

THEOREM 3.2. *Suppose* $\mathbf{g}(a)$ *is* ϕ_p*-optimal and* $\mathbf{g}(b)$ *is* ϕ_q*-optimal, where* $p, q \in (-\infty, 1)$. *If* $a < b$, *then* $p < q$.

PROOF. By (3.3) and (3.4), we have

$$[a/g(a)]^{p-1} \le -(n-1)g'(a_+) \tag{3.5}$$

and

$$[b/g(b)]^{q-1} \ge -(n-1)g'(b_-). \tag{3.6}$$

We shall show that $p \ge q$ would lead to a contradiction. By (2.4) and (2.5), $1 \le a/g(a) < b/g(b)$. If $p \ge q$, then $[b/g(b)]^{q-1} < [a/g(a)]^{p-1}$. On the other hand, $-(n-1)g'(b_-) \ge -(n-1)g'(a_+)$. These two inequalities together with (3.5) imply that $[b/g(b)]^{q-1} < -(n-1)g'(b_-)$, contradicting (3.6). □

By (2.5), for $a \in (a_L, a_R^S)$, (3.3) and (3.4) are the same as

$$\frac{\ln[-(n-1)g'(a_-)]}{\ln[a/g(a)]} + 1 \le p \le \frac{\ln[-(n-1)g'(a_+)]}{\ln[a/g(a)]} + 1. \tag{3.7}$$

For each $a \in (a_L, a_R^S)$, let

$$s(a) = \frac{\ln[-(n-1)g'(a_-)]}{\ln[a/g(a)]} + 1 \quad \text{and}$$

$$t(a) = \frac{\ln[-(n-1)g'(a_+)]}{\ln[a/g(a)]} + 1.$$

Then, for any $a \in (a_L, a_R^S)$, $\mathbf{g}(a)$ is ϕ_p-optimal for all $p \in [s(a), t(a)]$. Let

$$s(a_R^S) = \frac{\ln[-(n-1)g'(a_{R-}^S)]}{\ln[a_R^S/g(a_R^S)]} + 1 \quad \text{and} \quad t(a_R^S) = 1.$$

Then $\mathbf{g}(a_R^S)$ is ϕ_p-optimal for $p \in [s(a_R^S), t(a_R^S)]$. Furthermore, let $s(a_L) = -\infty$, and define

$$t(a_L) = \begin{cases} \dfrac{\ln[-(n-1)g'(a_{L+})]}{\ln[a_L/g(a_L)]} + 1, & \text{if } a_L \neq g(a_L), \\ -\infty, & \text{if } a_L = g(a_L). \end{cases}$$

Then $\mathbf{g}(a_L)$ is ϕ_p-optimal for $p \in [s(a_L), t(a_L)]$. Notice that when $a_L = g(a_L)$, (3.3) cannot be satisfied by any $p \in (-\infty, 1)$. In this case, $\mathbf{g}(a_L)$ is only $\phi_{-\infty}$-optimal. By Theorem 3.2, $t(a) < s(b)$ for any $a < b$; so $\{[s(a), t(a)]\}_{a \in [a_L, a_R^S]}$ gives a partition of $[-\infty, 1]$. If g is differentiable at a, then $s(a) = t(a)$, and $\mathbf{g}(a)$ is ϕ_p-optimal with respect to a single p; otherwise, it is ϕ_p-optimal with respect to all the p-values in an interval.

Consider once again the one-to-one correspondence (2.9) between $\mathscr{M}_S$ and $\mathscr{R}_S$. We see that any design $\varepsilon\xi_{\langle \mathbf{x}_1 \rangle} + (1-\varepsilon)\xi_{\langle \mathbf{x}_2 \rangle}$ in $\mathscr{M}_S$ is ϕ_p-optimal over Ξ with respect to all $p \in [s(a), t(a)]$, where $a = n^{-1}[\varepsilon\|\mathbf{x}_1\|_1^2 + (1-\varepsilon)\|\mathbf{x}_2\|_1^2]$. In particular, for any $\mathbf{x} \in E_S$, if g is differentiable at $a = n^{-1}\|\mathbf{x}\|_1^2$, then $\xi_{\langle \mathbf{x} \rangle}$ is ϕ_p-optimal for exactly one p-value, $p = s(a)$. Otherwise, it is ϕ_p-optimal for all p-values in the interval $[s(a), t(a)]$. On the other hand, for any two neighboring points $\mathbf{x}_1$ and $\mathbf{x}_2$ in E_S and $0 < \varepsilon < 1$, since $\mathbf{g}(a)$ lies in the interior of a line segment, g must be differentiable at $a = n^{-1}[\varepsilon\|\mathbf{x}_1\|_1^2 + (1-\varepsilon)\|\mathbf{x}_2\|_1^2]$. Therefore, $\varepsilon\xi_{\langle \mathbf{x}_1 \rangle} + (1-\varepsilon)\xi_{\langle \mathbf{x}_2 \rangle}$ is ϕ_p-optimal for exactly one p-value. As one moves from $\xi_{\langle \mathbf{x}_L \rangle}$ to $\xi_{\langle \mathbf{x}_R^S \rangle}$ in increasing order of $\|\mathbf{x}\|_1$, while connecting two neighboring designs $\xi_{\langle \mathbf{x} \rangle}$ with their mixtures, the points on $\mathscr{R}_S$ from $\mathbf{g}(a_L)$ to $\mathbf{g}(a_R^S)$ are generated in increasing order of a. In the meantime, ϕ_p-optimal designs are produced in increasing order of p.

4. Examples.

EXAMPLE 1. Consider the case where $X = \{\mathbf{x}: \sum_{i=1}^n x_i \le z$, and $x_i \ge 0$ for all $i\}$. Then $\tilde{X}$ consists of the n points: $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$, where $\mathbf{e}_i$ has ith entry equal to z and all the other entries are zero. Since $\|\mathbf{e}_i\|_1 = \|\mathbf{e}_i\|_2$ for all i, the condition in Theorem 2.3 is satisfied. Therefore the two orderings have the same minimally complete classes. Indeed, we have $\langle \mathbf{x}_R \rangle = \langle \mathbf{x}_L \rangle = \langle \mathbf{e}_i \rangle$. So

$\mathscr{R}_K$ $(=\mathscr{R}_S)$ contains only one point, and the essentially complete class $\mathscr{M}_K$ $(=\mathscr{M}_S)$ consists of one single design: $\xi_{\langle \mathbf{e}_1 \rangle}$, which puts equal weights on the n vertices $\mathbf{e}_1, \dots, \mathbf{e}_n$. This is the best design under both Kiefer and Schur orderings; in particular, it is ϕ_p-optimal for all $-\infty \le p \le 1$. We point out that it is also *universally optimal* in the sense of Kiefer (1975).

EXAMPLE 2. Let $X = \{\mathbf{x}: \sum_{i=1}^n x_i^2 \le r^2, \text{ and } x_i \ge 0 \text{ for all } i\}$. Then

$$\tilde{X} = \left\{\mathbf{x}: \sum_{i=1}^n x_i^2 = r^2, \text{ and } x_i \ge 0 \text{ for all } i\right\}$$

is the spherical surface. All the points in $\tilde{X}$ have the same L_2-norm, but may have different L_1-norms; so in this case we have different minimally complete classes for Kiefer and Schur orderings. Since all the points in $\tilde{X}$ have the same L_2-norm, $\mathbf{x}_L$ (respectively, $\mathbf{x}_R$) is obtained by minimizing (respectively, maximizing) $\|\mathbf{x}\|_1^2$ over $\mathbf{x} \in \tilde{X}$. Therefore $\mathbf{x}_R$ is the point with all entries equal to $r/\sqrt{n}$, and $\mathbf{x}_L$ can be any $\mathbf{e}_i$ whose ith entry is r, and all the other entries are 0, $1 \le i \le n$. For any a with $n^{-1}r^2 \le a \le r^2$, the set C_a described in the paragraph preceding Theorem 2.1 is the cross section $\tilde{X} \cap \{\mathbf{x}: \sum_{i=1}^n x_i = \sqrt{na}\}$, and the set E is obtained by picking an arbitrary point $\mathbf{x}_a$ from C_a, $n^{-1}r^2 \le a \le r^2$. Since $n^{-1}\|\mathbf{x}_a\|_1^2 = a$ and $(\|\mathbf{x}_a\|_2^2 - n^{-1}\|\mathbf{x}_a\|_1^2)/(n-1) = (r^2 - a)/(n-1)$, all the points $(n^{-1}\|\mathbf{x}_a\|_1^2, (\|\mathbf{x}_a\|_2^2 - n^{-1}\|\mathbf{x}_a\|_1^2)/(n-1))$, $n^{-1}r^2 \le a \le r^2$, are on a *straight line*. Hence $E_K = \{\mathbf{x}_L, \mathbf{x}_R\}$, and an essentially complete class for the Kiefer ordering consists of all the designs of the form $\varepsilon\xi_{\langle \mathbf{x}_L \rangle} + (1-\varepsilon)\xi_{\langle \mathbf{x}_R \rangle}$, $0 \le \varepsilon \le 1$. On the other hand, $\mathbf{x}_R^S = \mathbf{x}_L$ since $\mathbf{x}_L$ minimizes $\|\mathbf{x}\|_1^2$ over $\tilde{X}$. Therefore the essentially complete class $\mathscr{M}_S$ for the Schur ordering consists of one single design: $\xi_{\langle \mathbf{x}_L \rangle}$. As in Example 1, this design is universally optimal and is ϕ_p-optimal for all $-\infty \le p \le 1$.

It can easily be seen that $\varepsilon\xi_{\langle \mathbf{x}_L \rangle} + (1-\varepsilon)\xi_{\langle \mathbf{x}_R \rangle}$ has the same moment matrix as the uniform measure on C_a with $a = n^{-1}[\varepsilon r^2 + (1-\varepsilon)nr^2]$. Therefore an alternative essentially complete class for the Kiefer ordering is $\{\xi_a: n^{-1}r^2 \le a \le r^2\}$, where ξ_a is the uniform measure on C_a.

EXAMPLE 3. Consider $X = [0,1]^n$, which was treated in Cheng (1987) and Pukelsheim (1989). Then $\tilde{X}$ consists of all the vertices of X that have at least m entries equal to 1, where m is the integral part of $(n-1)/2$. Let $\mathbf{v}_k$ be a vertex with k entries equal to 1. Then

$$\|\mathbf{v}_k\|_1 = k \quad \text{and} \quad \|\mathbf{v}_k\|_2 = \sqrt{k}. \tag{4.1}$$

The condition in Theorem 2.3 is satisfied, so we have the same minimally complete class for the two orderings. In this case, $\mathbf{x}_L$ and $\mathbf{x}_R$ $(= \mathbf{x}_R^S)$ are obtained by maximizing $\|\mathbf{v}_k\|_2^2 - n^{-1}\|\mathbf{v}_k\|_1^2$ and $\|\mathbf{v}_k\|_1^2$, respectively, over $m \le k \le n$. It is easy to see that $\mathbf{x}_R (= \mathbf{x}_R^S) = \mathbf{v}_n$ and $\mathbf{x}_L = \mathbf{v}_m$. Then the set E is $\{\mathbf{v}_k: m \le k \le n\}$. By (4.1), the convex hull of $\{(n^{-1}\|\mathbf{x}\|_1^2, (\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2)/(n-1)): \mathbf{x} \in E\}$ is a polytope on the vertices $(k^2/n, (k - n^{-1}k^2)/(n-1))$, $m \le k \le n$. It follows that $E_K = E$; therefore the essentially complete class $\mathscr{M}_K$ $(=\mathscr{M}_S)$

consists of all the designs of the form $\varepsilon\xi_{\langle \mathbf{v}_k \rangle} + (1-\varepsilon)\xi_{\langle \mathbf{v}_{k+1} \rangle}$, $0 \le \varepsilon \le 1$, $m \le k \le n-1$, where $\xi_{\langle \mathbf{v}_k \rangle}$ is the uniform measure on all the vertices with k entries equal to 1. Cheng (1987) showed that each ξ_k is ϕ_p-optimal for the p-values in an interval, while each proper mixture $\varepsilon\xi_{\langle \mathbf{v}_k \rangle} + (1-\varepsilon)\xi_{\langle \mathbf{v}_{k+1} \rangle}$ is ϕ_p-optimal for one single p. This is because on the boundary of the polytope, the slopes change at the vertices only, which correspond to the designs $\xi_{\langle \mathbf{v}_k \rangle}$, $m < k < n$. Explicit formulas for ϕ_p-optimal designs as derived in Cheng (1987) can be obtained from (3.7), and there is no need to use the equivalence theorem. See also Pukelsheim [(1993), Section 14.10].

In Example 3, the function g is piecewise linear. We shall end the paper with an example in which g is differentiable.

EXAMPLE 4. Let X be the ball $\{\mathbf{x}: \sum_{i=1}^n (x_i - z)^2 \le r^2\}$, where $r \le z$. Then $X \subset \mathbb{R}_+^n$, and $\tilde{X} = \{\mathbf{x}: \sum_{i=1}^n (x_i - z)^2 = r^2$ and $\sum_{i=1}^n x_i \ge nz\}$. Again, the condition in Theorem 2.3 is satisfied, and the minimally complete classes for the two orderings coincide. It is clear that $\mathbf{x}_R$ is the point with all entries equal to $z + r/\sqrt{n}$. We notice that, for any $\mathbf{x}$ on the spherical surface $\{\mathbf{x}: \sum_{i=1}^n (x_i - z)^2 = r^2\}$, if $n^{-1}\|\mathbf{x}\|_1^2 = a$, then $\|\mathbf{x}\|_2^2 = r^2 - nz^2 + 2z\sqrt{na}$; so $\|\mathbf{x}\|_2^2 - n^{-1}\|\mathbf{x}\|_1^2 = r^2 - nz^2 + 2z\sqrt{na} - a$ is a function of a. Therefore $\mathbf{x}_L$ can be obtained by maximizing $r^2 - nz^2 + 2z\sqrt{na} - a$ with respect to a. The solution is $a = nz^2$. So $\mathbf{x}_L$ can be any point in the intersection of the spherical surface $\{\mathbf{x}: \sum_{i=1}^n (x_i - z)^2 = r^2\}$ and the hyperplane $\{\mathbf{x}: \sum_{i=1}^n x_i = nz\}$ that passes through the center $(z, z, \dots, z)^T$ of X. For any a with $nz^2 \le a \le n(z + r/\sqrt{n})^2$, the set C_a is the cross section $C_a = \{\mathbf{x}: \sum_{i=1}^n (x_i - z)^2 = r^2\} \cap \{\mathbf{x}: \sum_{i=1}^n x_i = \sqrt{na}\}$. Pick an arbitrary point $\mathbf{x}_a$ from C_a. Then the set E is $\{\mathbf{x}_a: nz^2 \le a \le n(z + r/\sqrt{n})^2\}$. Let

$$F = \left\{\left(a, (r^2 - nz^2 + 2z\sqrt{na} - a)/(n-1)\right): nz^2 \le a \le n\left(z + r/\sqrt{n}\right)^2\right\}.$$

Since $r^2 - nz^2 + 2z\sqrt{na} - a$ is a strictly concave function of a, $\mathscr{R}_K$, the upper boundary of the convex hull of F, is F itself. Therefore $E_K = E$. Furthermore, E contains no neighboring points because it is a continuum. It follows that the essentially complete class $\mathscr{M}_K$ $(=\mathscr{M}_S)$ consists of all the measures $\xi_{\langle \mathbf{x}_a \rangle}$, $nz^2 \le a \le n(z + r/\sqrt{n})^2$.

It is easy to see that $\xi_{\langle \mathbf{x}_a \rangle}$ has the same moment matrix as the uniform measure on C_a. Let this design be denoted ξ_a. Then an alternative essentially complete class for both Kiefer and Schur orderings is $\{\xi_a: nz^2 \le a \le n(z + r/\sqrt{n})^2\}$. For any such ξ_a, $g(a) = (r^2 - nz^2 + 2z\sqrt{na} - a)/(n-1)$. Since g is differentiable, each ξ_a, $nz^2 < a < n(z + r/\sqrt{n})^2$, is ϕ_p-optimal for a single p:

$$\begin{aligned} p &= \frac{\ln[-(n-1)g'(a)]}{\ln[a/g(a)]} + 1 \\ &= \frac{\ln\left[1 - z\sqrt{n/a}\right]}{\ln\left[(n-1)a/(r^2 - nz^2 + 2z\sqrt{na} - a)\right]} + 1. \end{aligned}$$

On the other hand, ξ_{nz^2} is $\phi_{-\infty}$-optimal, and $\xi_{n(z+r/\sqrt{n})^2}$ is ϕ_p-optimal for all

$$p \in \left[\frac{\ln[r\sqrt{n}/(nz + r\sqrt{n})]}{\ln\left[(n-1)n(z + r/\sqrt{n})^2/\left(r^2 - nz^2 + 2nz(z + r/\sqrt{n}) - n(z + r/\sqrt{n})^2\right)\right]} + 1, 1 \right].$$

Acknowledgment. I would like to thank a referee for helpful comments that improved the presentation of this article.

REFERENCES

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

HARWIT, M. and SLOANE, N. J. A. (1976). Masks for Hadamard transform optics, and weighing designs. *Applied Optics* **15** 107–114.

KIEFER, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Designs and Linear Models* (J. N. Srivastava, ed.) 333–353. North-Holland, Amsterdam.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

PUKELSHEIM, F. (1989). Complete class results for linear regression designs over the multidimensional cube. In *Contributions to Probability and Statistics. Essays in Honor of Ingram Olkin* (L. J. Gleser, M. D. Perlman, S. J. Press and A. R. Sampson, eds.) 349–356. Springer, New York.

PUKELSHEIM, F. (1993). *Optimal Design of Experiments.* Wiley, New York.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720