On third order rotatability

Norman R. Draper, Friedrich Pukelsheim

On Third Order Rotatability

By N.R. Draper[1] and F. Pukelsheim[2]

Abstract: Third order rotatability of experimental designs, moment matrices and information surfaces is investigated, using a Kronecker power representation. This representation complicates the model but greatly simplifies the theoretical development, and throws light on difficulties experienced in some previous work. Third order rotatability is shown to be characterized by the finitely many transformations consisting of permutations and a bi-axial 45 degree rotation, and the space of rotatable third order symmetric matrices is shown to be 20, independent of the number of factors m. A general Moore-Penrose inverse of a third order rotatable moment matrix is provided, leading to the information surface, and the corresponding optimality results are discussed. After a brief literature review, extensions to higher order models, the connections with tensor representations of classic matrix groups, and the evaluation of a general dimension formula, are all explored.

---

[1] Norman R. Draper, Department of Statistics, University of Wisconsin-Madison, Madison, WI 53706, USA. E-mail: `Draper@Stat.Wisc.Edu`

[2] Friedrich Pukelsheim, Institut für Mathematik, Universität Augsburg, Universitätstraße 8, DW-8900 Augsburg, Germany. E-mail: `Pukelsheim@Uni-Augsburg.De`

1   Introduction

In this paper we study third order rotatability of experimental designs, moment matrices and information surfaces.   The paper thus extends the second order rotatability results of Draper, Gaffke, and Pukelsheim (1991), henceforth quoted as DGP; see also Chapter 15 of Pukelsheim (1993).   Third order rotatability is of less practical relevance, but is very helpful in delineating the general underlying principles.

We again discuss the classical linear model of uncorrelated homoscedastic observations for fitting a polynomial model, here taken to be of order three.   For instance, for three input factors t_1, t_2, t_3 the true response surface η is given by a cubic polynomial in t_1, t_2, t_3,

$$
\begin{aligned}
\eta(t, \theta) = \theta_0 &+ t_1\theta_1 + t_2\theta_2 + t_3\theta_3 \\
&+ t_1^2\theta_{11} + t_1t_2\theta_{12} + t_1t_3\theta_{13} + t_2^2\theta_{22} + t_2t_3\theta_{23} + t_3^2\theta_{33} \\
&+ t_1^3\theta_{111} + t_1^2t_2\theta_{112} + t_1^2t_3\theta_{113} + t_1t_2^2\theta_{122} + t_1t_2t_3\theta_{123} \\
&+ t_1t_3^2\theta_{133} + t_2^3\theta_{222} + t_2^2t_3\theta_{223} + t_2t_3^2\theta_{233} + t_3^3\theta_{333}.
\end{aligned} \tag{1.1}
$$

Generally we admit $m \geq 3$ factors t_i so that (1.1) becomes a third order polynomial in $t_1, \ldots, t_m$, with coefficient vector θ to be estimated.

We define the third order regression vector for an m factor input vector $t = (t_1, \ldots, t_m)'$ by

$$
f(t) = \begin{pmatrix} 1 \\ t \\ t^{\otimes 2} \\ t^{\otimes 3} \end{pmatrix} \in \mathbb{R}^k, \qquad k = 1 + m + m^2 + m^3. \tag{1.2}
$$

The Kronecker powers $t^{\otimes 2} = t \otimes t$ and $t^{\otimes 3} = t \otimes t \otimes t$ repeat the individual mixed second order terms twice, and the third order terms either six or three times depending on the number of different subscripts, ijk or iij. Thus (1.1) is augmented by additional terms $\theta_{21}, \theta_{31}, \theta_{32}$, etc.   The very same point is familiar from treating dispersion matrices as matrices, and *not* as arrays of a minimal number of functionally independent terms.   Also, the overparametrization points towards interesting connections with tensor representations and produces a simple calculus.

For instance, if t is rotated into Rt where R is a matrix in the group $\mathrm{Orth}(m)$ of orthogonal $m \times m$ matrices, then the vector $f(t)$ undergoes the transformation $Q_R f(t)$

where the $k \times k$ matrix Q_R is given by

$$Q_R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & R & 0 & 0 \\ 0 & 0 & R^{\otimes 2} & 0 \\ 0 & 0 & 0 & R^{\otimes 3} \end{pmatrix}. \tag{1.3}$$

As R varies over $\mathrm{Orth}(m)$ the matrices Q_R form a subgroup $\mathcal{Q}$ of $\mathrm{Orth}(k)$. We let the group $\mathcal{Q}$ act by congruence, $M \mapsto Q_R M Q_R'$, and study invariance of symmetric matrices (Section 2) and of moment matrices (Section 3).

The Kronecker power representation in (1.2) makes the induced matrix Q_R in (1.3) orthogonal. Hence the congruence transformation $Q_R M Q_R'$ preserves eigenvalues, quite similar to the Schläflian calculus of Box and Hunter (1957), Draper (1984), or Arap Koske and Patil (1989). In contrast to (1.2), Kiefer and Galil (1977, page 30) use a minimal system of squares and cross products for which the induced matrix (C_Q in their notation) need not be orthogonal. It is this lack of orthogonality that, in the setting of Kiefer and Galil (1977, 1979), Heiligers and Schneider (1992), and Gaffke and Heiligers (1992), makes it possible for non-rotatable moment matrices to be A- or E-optimal. It is for the same reason that the rotatability measure of Khuri (1988, 1992) is not invariant under design rotation, as pointed out by Draper and Pukelsheim (1990).

These difficulties do not appear with the Kronecker power representation (1.2). Then a projection of an arbitrary moment matrix onto the rotatable moment matrices improves any concave optimality criterion ϕ which depends on the moment matrix only through its eigenvalues. Hence ϕ-optimality among *rotatable* moment matrices implies ϕ-optimality among *all* moment matrices, and the above mentioned restriction to rotatable designs ceases to be a true restriction. For second order models the ensuing complete class theorem is given in Pukelsheim (1993, Theorem 15.19). For third order models we establish an appropriate result in Section 4.

The paper is organized as follows. Our main result, in Section 2, is that third order rotatability is characterized by the finitely many transformations consisting of permutations and a bi-axial 45° rotation. We show that the space of rotatable symmetric matrices has dimension 20, independent of m, and provide a basis for it. In Section 3 we calculate the Moore–Penrose inverse of rotatable moment matrices. Section 4 contains the corresponding optimality discussion. Section 5 reviews the literature on third order rotatability. Section 6 concludes the paper with some general remarks on extensions to higher order models, interrelations with tensor representations of classical matrix groups, and a dimension formula.

2   Third Order Rotatability of Symmetric Matrices

Before presenting Theorem 2.1 we make some preparatory comments. Part (iii) of
the theorem states that the space of invariant symmetric matrices has dimension 20,
irrespective of the value of m. The dimension formula for a general order $d \geq 4$ is
outlined in Section 6. We use $k = 1 + m + m^2 + m^3$ from (1.2), with $m \geq 3$.

The result parallels Theorem 6.1 of DGP. There, in part (ii), we used sign change
matrices, permutation matrices, the orthodiagonal reflection and, for $m = 2$, a 45°
rotation. It transpires that the orthodiagonal reflection can be replaced by a bi-axial
45° rotation throughout; similarly the sign changes can be dispensed with since they
can be generated by permutations and any bi-axial 45° rotation. Part (ii) below is
presented in this more streamlined form.

For a concise parametrization of the invariant symmetric matrices in part (ii) we
define
$$e_{ij} = e_i \otimes e_j \in \mathbb{R}^{m^2}, \qquad e_{ijk} = e_i \otimes e_j \otimes e_k \in \mathbb{R}^{m^3},$$
where e_i is the ith Euclidean unit vector of $\mathbb{R}^m$ and $i, j, k = 1, \ldots, m$.

*Theorem 2.1: Let M be a symmetric $k \times k$ matrix, not necessarily a moment matrix
of a design. Then the following three statements are equivalent:*

(i) *M is third order rotatable, that is, $M = Q_R M Q'_R$ for all $R \in \mathrm{Orth}(m)$.*

(ii) *M is invariant under the finitely many matrices Q_R where R is any permutation
matrix or a bi-axial 45° rotation.*

(iii) *There exist 20 scalars $\alpha, \beta, \gamma, \delta_1, \delta_2, \delta_3, \eta_1, \eta_2, \eta_3, \epsilon_1, \ldots, \epsilon_{11} \in \mathbb{R}$ such that M has
the form*
$$M = \begin{pmatrix} \alpha & 0 & \gamma \sum_i e'_{ii} & 0 \\ 0 & \beta \sum_i e_i e'_i & 0 & G(\eta)' \\ \gamma \sum_i e_{ii} & 0 & F(\delta) & 0 \\ 0 & G(\eta) & 0 & S(\epsilon) \end{pmatrix} \tag{2.1}$$

where

$$F(\delta) = \delta_1 \sum_{i,j} e_{ij} e'_{ij} + \delta_2 \sum_{i,j} e_{ji} e'_{ij} + \delta_3 \sum_{i,j} e_{ii} e'_{jj}, \tag{2.2}$$

$$G(\eta) = \eta_1 \sum_{i,j} e_{iji} e'_j + \eta_2 \sum_{i,j} e_{jii} e'_j + \eta_3 \sum_{i,j} e_{iij} e'_j, \tag{2.3}$$

$$S(\epsilon) = \epsilon_1 \sum_{i,j,k} e_{ijk} e'_{ijk} + \epsilon_2 \sum_{i,j,k} e_{ikj} e'_{ijk} + \epsilon_3 \sum_{i,j,k} e_{jik} e'_{ijk} + \epsilon_4 \sum_{i,j,k} e_{kji} e'_{ijk}$$

$$+ \epsilon_5 \sum_{i,j,k} (e_{jki}e'_{ijk} + e_{kij}e'_{ijk})$$

$$+ \epsilon_6 \sum_{i,j,k} e_{iji}e'_{kjk} + \epsilon_7 \sum_{i,j,k} e_{jii}e'_{jkk} + \epsilon_8 \sum_{i,j,k} e_{iij}e'_{kkj}$$

$$+ \epsilon_9 \sum_{i,j,k} (e_{iji}e'_{jkk} + e_{jii}e'_{kjk}) + \epsilon_{10} \sum_{i,j,k} (e_{iji}e'_{kkj} + e_{iij}e'_{kjk})$$

$$+ \epsilon_{11} \sum_{i,j,k} (e_{jii}e'_{kkj} + e_{iij}e'_{jkk}). \tag{2.4}$$

Proof: Part (i) implies (ii), since (ii) comprises fewer transformations than (i). Part (iii) implies (i) by direct verification. To illustrate the type of calculations involved let $P = \sum_{i,j,k} e_{kij}e'_{ijk}$ be the second matrix that appears with ϵ_5. For $x, y, z \in \mathbb{R}^m$ the action of P is

$$P(x \otimes y \otimes z) = \sum_{i,j,k} x_i y_j z_k \, e_k \otimes e_i \otimes e_j = z \otimes x \otimes y,$$

$$R^{\otimes 3} P(x \otimes y \otimes z) = Rz \otimes Rx \otimes Ry = P(Rx \otimes Ry \otimes Rz) = PR^{\otimes 3}(x \otimes y \otimes z).$$

It follows that $R^{\otimes 3}$ and P commute. Hence P is invariant, $R^{\otimes 3} P R^{\otimes 3\prime} = P$. This type of argument covers the first six matrices in $S(\epsilon)$, attached to the subscripts $1, 2, 3, 4, 5, 5$ of ϵ. A variation of the argument applies to the other nine terms. As a generic case we consider the matrix $Q = \sum_{i,j,k} e_{iij}e'_{kkj}$ of ϵ_8. For $x, y, z \in \mathbb{R}^m$ its action is

$$Q(x \otimes y \otimes z) = \sum_{i,j,k} x_k y_k z_j \, e_i \otimes e_i \otimes e_j = x'y \, (\text{vec} \, I_m) \otimes z,$$

$$R^{\otimes 3} Q(x \otimes y \otimes z) = x'y \, (\text{vec} \, I_m) \otimes Rz = Q(Rx \otimes Ry \otimes Rz) = QR^{\otimes 3}(x \otimes y \otimes z),$$

where we have used $R^{\otimes 2}(\text{vec} \, I_m) = \text{vec}(RR') = \text{vec} \, I_m$ from DGP, page 144. Again $R^{\otimes 3}$ and Q commute, that is, Q is invariant.

The major part of the proof shows that part (ii) implies (iii). Let M be a symmetric matrix which satisfies (ii). By assumption, we are provided with all permutations and a single bi-axial $45°$ rotation. Multiple combinations of these transformations yield all other bi-axial $45°$ rotations, as well as all sign changes. Thus Theorem 6.1 of DGP gives (2.1) except for $G(\eta)$ and $S(\epsilon)$ which arise from the introduction of third order terms.

In order to investigate in M the $m^3 \times m$ submatrix $G = \sum_{i,j,k,l} g_{ijk,l} \, e_{ijk}e'_l$, the m^4 numbers $g_{ijk,l}$ are rearranged into the $m^2 \times m^2$ matrix $D = \sum_{i,j,k,l} g_{ijk,l} \, e_{ij}e'_{kl}$

which is not necessarily symmetric. It is evident that third order rotatability may be expressed in either terms,

$$G = R^{\otimes 3} G R' \qquad \Longleftrightarrow \qquad D = R^{\otimes 2} D R^{\otimes 2\prime}. \tag{2.5}$$

The latter entails $D = F(\eta)$ since the arguments of DGP, page 145f, carry over despite the apparent initial lack of symmetry of D in (2.5). This forces G to attain the form $G(\eta)$ of (2.3).

Finally we turn to the $m^3 \times m^3$ bottom right block S in M,

$$S = \Big(\sum_{i,j,k} e_{ijk} e'_{ijk}\Big) S \Big(\sum_{a,b,c} e_{abc} e'_{abc}\Big) = \sum_{i,j,k} \sum_{a,b,c} s_{ijk,abc}\, e_{ijk} e'_{abc}. \tag{2.6}$$

Invariance under any sign change matrix $R = \mathrm{diag}(\varepsilon_1, \ldots, \varepsilon_m)$, with $\varepsilon_i = \pm 1$, yields an identity for the entries of S,

$$s_{ijk,abc} = e'_{ijk} S e_{abc} = e'_{ijk} R^{\otimes 3} S R^{\otimes 3} e_{abc} = \varepsilon_i \varepsilon_j \varepsilon_k \varepsilon_a \varepsilon_b \varepsilon_c\, s_{ijk,abc}. \tag{2.7}$$

Hence these entries vanish when the number of distinct subscripts is 6, or 5, or 4, or 3 with at least one odd multiplicity, or 2 with at least one odd multiplicity.

This leaves 3 distinct subscripts of multiplicity two each, or 2 distinct subscripts of multiplicities two and four, or 1 subscript of multiplicity six. This and the symmetry of S entail

$$
\begin{aligned}
S = \sum_{i \neq j \neq k \neq i} \Big(& s_{ijk,ijk}\, e_{ijk} e'_{ijk} + s_{ikj,ijk}\, e_{ikj} e'_{ijk} + s_{jik,ijk}\, e_{jik} e'_{ijk} + s_{kji,ijk}\, e_{kji} e'_{ijk} \\
& + s_{jki,ijk} \left(e_{jki} e'_{ijk} + e_{ijk} e'_{jki} \right) \\
& + s_{iji,kjk}\, e_{iji} e'_{kjk} + s_{jii,jkk}\, e_{jii} e'_{jkk} + s_{iij,kkj}\, e_{iij} e'_{kkj} \\
& + s_{iji,jkk} \left(e_{iji} e'_{jkk} + e_{jkk} e'_{iji} \right) + s_{iji,kkj} \left(e_{iji} e'_{kkj} + e_{kkj} e'_{iji} \right) \\
& + s_{jii,kkj} \left(e_{jii} e'_{kkj} + e_{kkj} e'_{jii} \right) \Big) \\
+ \sum_{i \neq j} \Big(& s_{ijj,ijj}\, e_{ijj} e'_{ijj} + s_{jij,jij}\, e_{jij} e'_{jij} + s_{jji,jji}\, e_{jji} e'_{jji} \\
& + s_{iij,jjj} \left(e_{iij} e'_{jjj} + e_{jjj} e'_{iij} \right) + s_{ijj,jij} \left(e_{ijj} e'_{jij} + e_{jij} e'_{ijj} \right) \\
& + s_{iji,jjj} \left(e_{iji} e'_{jjj} + e_{jjj} e'_{iji} \right) + s_{ijj,jji} \left(e_{ijj} e'_{jji} + e_{jji} e'_{ijj} \right) \\
& + s_{jii,jjj} \left(e_{jii} e'_{jjj} + e_{jjj} e'_{jii} \right) + s_{jij,jji} \left(e_{jij} e'_{jji} + e_{jji} e'_{jij} \right) \Big) \\
+ \sum_i & s_{iii,iii}\, e_{iii} e'_{iii}.
\end{aligned} \tag{2.8}
$$

There remain $11m(m-1)(m-2) + 9m(m-1) + m = 11m^3 - 24m^2 + 14m$ different entries to be determined.

For a further reduction we take R to be any permutation matrix to obtain $s_{ijk,abc} = s_{\pi(i)\pi(j)\pi(k),\pi(a)\pi(b)\pi(c)}$, for every permutation π of $1,\ldots,m$. Now, such a permutation does not change the pattern of those of the six subscripts which were previously equal. Hence the $m(m-1)(m-2)$ coefficients $s_{ijk,ijk}$ are equal, all $s_{ikj,ijk}$ are equal, and so on. We denote the first 11 coefficients by $\epsilon_1,\ldots,\epsilon_{11}$, the next 9 by $\eta_1,\ldots,\eta_9$, and the last one by θ,

$$s_{ijk,ijk} = \epsilon_1, \qquad \ldots, \qquad s_{ijj,ijj} = \eta_1, \qquad \ldots, \qquad s_{iii,iii} = \theta. \qquad (2.9)$$

This reduces the number of distinct entries to $11 + 9 + 1 = 21$, independent of m.

In order to express $\eta_1,\ldots,\eta_9$ and θ as functions of $\epsilon_1,\ldots,\epsilon_{11}$ we apply the $45°$ rotation R_0 of the (t_1, t_2) plane,

$$R_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{2} \end{pmatrix}, \qquad (2.10)$$

and evaluate $x'Sy = x'R_0^{\otimes 3} S R_0^{\otimes 3\prime} y$. For $\eta_1 = s_{ijj,ijj}$ we take $x = e_{311}$ and $y = e_{311}$, $\ldots$, for $\eta_9 = s_{jij,jji}$ we take $x = e_{131}$ and $y = e_{113}$; for θ we take $x = y = e_{111}$:

$$\eta_1 = \epsilon_1 + \epsilon_2 + \epsilon_7, \qquad \eta_2 = \epsilon_1 + \epsilon_4 + \epsilon_6, \qquad \eta_3 = \epsilon_1 + \epsilon_3 + \epsilon_8,$$

$$\eta_4 = \epsilon_8 + \epsilon_{10} + \epsilon_{11}, \qquad \eta_5 = \epsilon_3 + \epsilon_5 + \epsilon_9, \qquad \eta_6 = \epsilon_6 + \epsilon_9 + \epsilon_{10}, \qquad (2.11)$$

$$\eta_7 = \epsilon_4 + \epsilon_5 + \epsilon_{11}, \qquad \eta_8 = \epsilon_7 + \epsilon_9 + \epsilon_{11}, \qquad \eta_9 = \epsilon_2 + \epsilon_5 + \epsilon_{10};$$

$$\theta = \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4 + 2\epsilon_5 + \epsilon_6 + \epsilon_7 + \epsilon_8 + 2\epsilon_9 + 2\epsilon_{10} + 2\epsilon_{11}. \qquad (2.12)$$

Thus S in (2.8) involves the coefficients $\epsilon_1,\ldots,\epsilon_{11}$ only. The three sums over distinct subscripts reduce to a single sum over unrestricted subscripts. Applying the unrestricted sum to the individual terms we find $\sum_{i,j,k} e_{ijk} e'_{jki} = \sum_{i,j,k} e_{kij} e'_{ijk}$, etc. This establishes the pattern $S(\epsilon)$ of (2.4). □

The matrices accompanying $\epsilon_1,\ldots,\epsilon_{11}$ in (2.4) can be shown to be linearly independent. Hence in general $S(\epsilon)$ cannot be reduced to fewer terms. This is in contrast to moment matrices which, due to their structure, permit a further simplification, discussed next.

3   Third Order Rotatability of Moment Matrices

A design τ on an experimental domain $\mathcal{T} \subseteq \mathbb{R}^m$ is a probability measure with finite support. The points $t \in \mathcal{T}$ which receive a positive weight under τ are the support points of τ and designate the experimental conditions under which the response is to be observed. The weight $\tau(t)$ gives the proportion of observations to be drawn under t. For our rotatability discussion we choose, as the rotatable experimental domain, the ball of radius $\sqrt{m}$,

$$\mathcal{T}_{\sqrt{m}} = \{t \in \mathbb{R}^m : \|t\| \leq \sqrt{m}\},$$

so that the vertices of the unit cube $\{\pm 1\}^m$ lie on the bounding sphere.

The third order moment matrix of a design τ on $\mathcal{T}_{\sqrt{m}}$ is defined to be

$$M(\tau) = \int_{\mathcal{T}_{\sqrt{m}}} f(t)f(t)' \, d\tau, \tag{3.1}$$

with regression function f given by (1.2). This is a symmetric and nonnegative definite $k \times k$ matrix. Rotatability of $M(\tau)$ entails the pattern (2.1): all odd moments vanish, and the even moments can be expressed through three moment parameters of τ defined by

$$\lambda_2(\tau) = \int_{\mathcal{T}_{\sqrt{m}}} t_1^2 \, d\tau, \qquad \lambda_4(\tau) = \int_{\mathcal{T}_{\sqrt{m}}} t_1^2 t_2^2 \, d\tau, \qquad \lambda_6(\tau) = \int_{\mathcal{T}_{\sqrt{m}}} t_1^2 t_2^2 t_3^2 \, d\tau.$$

The following theorem describes the structure of $M(\tau)$ more precisely, again using $k = 1 + m + m^2 + m^3$ from (1.2), with $m \geq 3$.

Theorem 3.1:   Let M be a symmetric $k \times k$ matrix. Then M is a third order rotatable moment matrix on the experimental domain $\mathcal{T}_{\sqrt{m}}$ if and only if there exist scalars λ_2, λ_4 and λ_6 fulfilling either (3.2), or (3.3), or (3.4):

$$\lambda_2 = 0 = \lambda_4 = \lambda_6; \tag{3.2}$$

$$\lambda_2 = 1, \qquad \lambda_4 = \frac{m}{m+2}, \qquad \lambda_6 = \frac{m^2}{(m+2)(m+4)}; \tag{3.3}$$

$$\lambda_2 \in (0, 1), \tag{3.4a}$$

$$\lambda_4 \in \left[\frac{m}{m+2}\lambda_2^2, \; \frac{m}{m+2}\lambda_2\right], \tag{3.4b}$$

$$\lambda_6 \in \left[\frac{m+2}{m+4}\frac{\lambda_4^2}{\lambda_2}, \; \frac{m+2}{m+4}\left(\frac{\lambda_4^2}{\lambda_2} + \frac{(\lambda_4 - \frac{m}{m+2}\lambda_2^2)(\frac{m}{m+2}\lambda_2 - \lambda_4)}{\lambda_2(1 - \lambda_2)}\right)\right]; \tag{3.4c}$$

such that M takes the form

$$M = \begin{pmatrix} 1 & 0 & \lambda_2 v'_m & 0 \\ 0 & \lambda_2 I_m & 0 & \lambda_4 G'_m \\ \lambda_2 v_m & 0 & \lambda_4 F_m & 0 \\ 0 & \lambda_4 G_m & 0 & \lambda_6 S_m \end{pmatrix} \tag{3.5}$$

where

$$v_m = \sum_i e_{ii}, \tag{3.6}$$

$$F_m = \sum_{i,j} e_{ij} e'_{ij} + \sum_{i,j} e_{ji} e'_{ij} + \sum_{i,j} e_{ii} e'_{jj}, \tag{3.7}$$

$$G_m = \sum_{i,j} e_{iji} e'_j + \sum_{i,j} e_{jii} e'_j + \sum_{i,j} e_{iij} e'_j, \tag{3.8}$$

$$\begin{aligned} S_m = &\sum_{i,j,k} e_{ijk} e'_{ijk} + \sum_{i,j,k} e_{ikj} e'_{ijk} + \sum_{i,j,k} e_{jik} e'_{ijk} + \sum_{i,j,k} e_{kji} e'_{ijk} \\ &+ \sum_{i,j,k} (e_{jki} e'_{ijk} + e_{kij} e'_{ijk}) \\ &+ \sum_{i,j,k} e_{iji} e'_{kjk} + \sum_{i,j,k} e_{jii} e'_{jkk} + \sum_{i,j,k} e_{iij} e'_{kkj} \\ &+ \sum_{i,j,k} (e_{iji} e'_{jkk} + e_{jii} e'_{kjk}) + \sum_{i,j,k} (e_{iji} e'_{kkj} + e_{iij} e'_{kjk}) \\ &+ \sum_{i,j,k} (e_{jii} e'_{kkj} + e_{iij} e'_{jkk}). \end{aligned} \tag{3.9}$$

Proof: It is convenient to work with the uniform distribution u_r on the sphere $\{t \in \mathbb{R}^m : \|t\| = r\}$ of radius r. For the lack of a finite support this is *not* a design as defined at the beginning of this section; however, any proper design which has the same moments up to order 6 as has u_r will do just as well. The moment parameters of u_r are known to be

$$\lambda_2(u_r) = \frac{r^2}{m}, \qquad \lambda_4(u_r) = \frac{r^4}{m(m+2)}, \qquad \lambda_6(u_r) = \frac{r^6}{m(m+2)(m+4)}.$$

For the direct part of the proof, let M be a moment matrix that is attained by the design τ. The rotatability pattern (2.1) entails (3.5), with $\lambda_j = \lambda_j(\tau)$ for $j = 2, 4, 6$. We need to verify the ranges (3.2–4). If $\lambda_2(\tau) = 0$, then $\tau = u_0$ is the one-point design in 0 and we get (3.2). If $\lambda_2(\tau) = 1$, then τ is supported by the sphere of radius $\sqrt{m}$ and, because of rotatability, it behaves like the uniform distribution $u_{\sqrt{m}}$, whence we get (3.3).

Otherwise (3.4a) applies. For $t \in \mathcal{T}_{\sqrt{m}}$ we introduce the squared norm and scale it to lie in the unit interval,

$$s(t) = \frac{1}{m}t't \in [0,1].$$

The moments of s under τ are readily found to be

$$\int s(t) \, d\tau = \lambda_2 = a,$$

$$\int s^2(t) \, d\tau = \frac{m+2}{m}\lambda_4 = b, \tag{3.10}$$

$$\int s^3(t) \, d\tau = \frac{(m+2)(m+4)}{m^2}\lambda_6 = c,$$

say. Evaluation of the determinants of appropriately (see discussion of (3.20)) chosen 2×2 moment matrices of functions of s yields bounds for b and c:

$$0 \le \det \int \begin{pmatrix} 1 & s(t) \\ s(t) & s^2(t) \end{pmatrix} d\tau = \det \begin{pmatrix} 1 & a \\ a & b \end{pmatrix} = b - a^2,$$

$$0 \le \int \left(s(t) - s^2(t)\right) d\tau = a - b,$$

$$0 \le \det \int \begin{pmatrix} s(t) & s^2(t) \\ s^2(t) & s^3(t) \end{pmatrix} d\tau = \det \begin{pmatrix} a & b \\ b & c \end{pmatrix} = ac - b^2, \tag{3.11}$$

$$0 \le \det \int \begin{pmatrix} 1 - s(t) & s(t) - s^2(t) \\ s(t) - s^2(t) & s^2(t) - s^3(t) \end{pmatrix} d\tau = \det \begin{pmatrix} 1 - a & a - b \\ a - b & b - c \end{pmatrix}$$
$$= (1-a)b - (a-b)^2 - (1-a)c.$$

That is, in the present case we obtain the following ranges for a, b, c,

$$a \in (0,1), \tag{3.12a}$$

$$b \in [a^2, a] = [a^2, a^2 + (a - 0)(1 - a)], \tag{3.12b}$$

$$c \in \left[\frac{b^2}{a}, \frac{b^2}{a} + \frac{(b - a^2)(a - b)}{(a - 0)(1 - a)}\right], \tag{3.12c}$$

which translate back into (3.4a–c). This proves the direct part. Note that the upper limit in (3.12b) depends on the position of a in the interval (3.12a), similarly the upper limit in (3.12c) involves the positions of both a and b within their respective intervals.

For the converse part, the matrix M of (3.5) evidently enjoys the rotatability pattern (2.1). We face the task of finding a design τ which has M for its moment matrix. In case of (3.2) we choose $\tau = u_0$. In case of (3.3) we use $\tau = u_{\sqrt{m}}$.

In case of (3.4) we conjecture that M is attained by the mixture of two spherical uniform distributions,

$$\tau_{\alpha,r,R} = \alpha u_{R\sqrt{m}} + (1 - \alpha)u_{r\sqrt{m}}, \tag{3.13}$$

with $\alpha \in [0,1]$ and $0 \leq r < R \leq 1$. Any design (3.13) is rotatable, with moment parameters $\lambda_j(\tau_{\alpha,r,R}) = \alpha\lambda_j(u_{R\sqrt{m}}) + (1-\alpha)\lambda_j(u_{r\sqrt{m}})$ for $j = 2, 4, 6$. We set these equal to λ_j of (3.4a–c) to obtain the system of equations

$$\alpha R^2 + (1-\alpha)r^2 = a,$$
$$\alpha R^4 + (1-\alpha)r^4 = b, \tag{3.14}$$
$$\alpha R^6 + (1-\alpha)r^6 = c,$$

where a, b, c from (3.10) range through the intervals in (3.12). The system (3.14) implies the identities

$$b - a^2 = \alpha(1-\alpha)(R^2 - r^2)^2,$$
$$c - ab = \alpha(1-\alpha)(R^2 - r^2)^2(R^2 + r^2), \tag{3.15}$$
$$ac - b^2 = \alpha(1-\alpha)(R^2 - r^2)^2 R^2 r^2.$$

If $b - a^2 > 0$ then R^2 and r^2 are the two zeros of the parabola

$$(x - R^2)(x - r^2) = x^2 - (R^2 + r^2)x + R^2 r^2 = x^2 - \frac{c - ab}{b - a^2}x + \frac{ac - b^2}{b - a^2}. \tag{3.16}$$

It follows that, if the system (3.14) is solvable in r^2, R^2 and α, then the solutions are uniquely given by

$$R^2 = \frac{c - ab + w}{2(b - a^2)}, \qquad r^2 = \frac{c - ab - w}{2(b - a^2)}, \qquad \alpha = \frac{a - r^2}{R^2 - r^2}, \tag{3.17}$$

where w is the positive root of

$$w^2 = (c - ab)^2 - 4(b - a^2)(ac - b^2). \tag{3.18}$$

It remains to verify our conjecture.

If $b = a^2$ then the range for c in (3.12c) collapses and we get $c = b^2/a = a^3$. Hence we realize M through a single uniform distribution $u_{r\sqrt{m}}$ with $r^2 = \lambda_2 \in (0,1)$.

If $b > a^2$ then we use the lower bound $ac \geq b^2$ in (3.12c) to obtain $c \geq ab$. With w^2 defined by (3.18) we find $(c - ab)^2 \geq w^2$, and

$$w^2 = 4(b - a^2)^3 + (2a^3 - 3ab + c)^2 > 0. \tag{3.19}$$

Hence R^2 and r^2 are indeed well defined by (3.17) and satisfy $0 \leq r^2 < R^2 \leq 1$, the inequality $R^2 \leq 1$ following from the upper bound in (3.12c). Moreover (3.19) yields $w > \pm(2a^3 - 3ab + c)$, and this converts into $r^2 < a < R^2$. Hence α, as defined by (3.17), lies in the open interval $(0,1)$. Straightforward but tedious calculation verifies the alternative representations

$$\alpha = \frac{b - r^4}{R^4 - r^4} = \frac{c - r^6}{R^6 - r^6}.$$

Therefore R^2, r^2 and α from (3.17) solve (3.14), whence $\tau_{\alpha,r,R}$ from (3.13) has M for its moment matrix. $\qquad\square$

With a, b, c from (3.10) conditions (3.2–4) hold if and only if both of the 2×2 matrices, which appear in (3.11),

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}, \qquad \begin{pmatrix} 1-a & a-b \\ a-b & b-c \end{pmatrix} \tag{3.20}$$

are nonnegative definite. That this is necessary and sufficient for $(1, a, b, c)$ to be a sequence of moments over the interval $[0, 1]$ is Theorem IV.1.1 of Karlin and Studden (1966). While the general moment theory thus provides the abstract existence of a measure attaining the given moments, we have actually constructed such a measure explicitly, for the third order model.

The construction of $\tau_{\alpha,r,R}$ together with the other cases also proves the following.

Corollary 3.2:   Every third order rotatable moment matrix can be realized by a single spherical uniform distribution, or by a nondegenerate mixture of two spherical uniform distributions. A nondegenerate mixture, that is $\tau_{\alpha,r,R}$ from (3.13) with $\alpha \in (0, 1)$ and $0 < r < R \le 1$, occurs if and only if the moments satisfy $(m+4)\lambda_2\lambda_6 > (m+2)\lambda_4^2$. □

This corollary improves upon Theorem 3.2.1 of Kiefer (1960), and Theorem X.7.4 of Karlin and Studden (1966), for third order models. Their results say that the D-optimal moment matrix is rotatable; using the Kiefer–Wolfowitz Equivalence Theorem they then deduce that the D-optimal design is a unique mixture of two spherical uniform distributions. Our result applies the rotatability properties directly, and says that *every* third order rotatable moment matrix, optimal or not, may be obtained from a mixture of two spherical uniform distributions (in the nondegenerate case; a single uniform distribution applies in the degenerate case). We return to the discussion of optimality properties in Section 4.

Since the Kronecker powers in (1.2) include a redundant repetition of higher order terms, any moment matrix $M(\tau)$ in (3.1) is necessarily rank deficient. Therefore we replace regular matrix inversion by Moore–Penrose inversion. For a rotatable moment matrix M from (3.5), we base the calculation of the Moore–Penrose inverse M^+ on the eigenvalue decomposition of M.

To this end we introduce the set $\mathcal{S}_2$ of permutations $\rho(ij)$ of the two symbols i, j, and the set $\mathcal{S}_3$ of the permutations $\sigma(ijk)$ of the three symbols i, j, k,

$$\mathcal{S}_2 = \{ij, ji\}, \qquad \mathcal{S}_3 = \{ijk, ikj, jik, kji, jki, kij\}.$$

These permutations visibly appear in the first two terms of the fourth order block F_m in (3.7), and the first six terms of the sixth order block S_m in (3.9). We define two auxiliary matrices,

$$\Pi_4 = \frac{1}{2} \sum_{i,j} \sum_{\rho \in \mathcal{S}_2} e_{\rho(ij)} e_{ij}', \qquad \Pi_6 = \frac{1}{6} \sum_{i,j,k} \sum_{\sigma \in \mathcal{S}_3} e_{\sigma(ijk)} e_{ijk}'. \tag{3.21}$$

The matrices Π_4 and Π_6 are understood best by their actions as linear transformations on $\mathbb{R}^{m^2}$ and $\mathbb{R}^{m^3}$, just as the matrices P and Q in the proof of Theorem 3.1, or the vec permutation matrix $I_{m,m}$ of (5.10) in DGP.

Given vectors $x, y, z \in \mathbb{R}^m$, we find that $\Pi_4(x \otimes y)$ is the average of $x \otimes y$ and $y \otimes x$, while $\Pi_6(x \otimes y \otimes z)$ is the average of $x \otimes y \otimes z, x \otimes z \otimes y, y \otimes x \otimes z, z \otimes y \otimes x, y \otimes z \otimes x, z \otimes x \otimes y$. In this sense Π_4 and Π_6 act as symmetrizers of 2- and 3-fold Kronecker products. The fact that they are averaging operators makes them symmetric and idempotent matrices, as is readily verified from the definition (3.21). Their general usefulness in calculating higher order moments becomes also evident in formula (2.3) of Pukelsheim (1980).

The matrices F_m from (3.7) and S_m from (3.9) now take a succinct form,

$$F_m = 2\Pi_4 + v_m v_m', \qquad S_m = 6\Pi_6 + G_m G_m'. \tag{3.22}$$

In view of $v_m' v_m = m$ and $G_m' G_m = 3(m+2)I_m$, we define the projection matrices

$$\widetilde{F}_m = \Pi_4 - \frac{1}{m} v_m v_m', \qquad \widetilde{S}_m = \Pi_6 - \frac{1}{3(m+2)} G_m G_m', \tag{3.23}$$

with associated degrees of freedom

$$\text{trace } \widetilde{F}_m = \frac{1}{2} m(m+1) - 1, \qquad \text{trace } \widetilde{S}_m = \frac{1}{6} m(m+1)(m+2) - m. \tag{3.24}$$

From $\widetilde{F}_m v_m = 0$ and $\widetilde{S}_m G_m = 0$ we get $F_m \widetilde{F}_m = 2\widetilde{F}_m$ and $S_m \widetilde{S}_m = 6\widetilde{S}_m$.

Hence the moment matrix M of (3.5) has eigenvalues $2\lambda_4$ and $6\lambda_6$, with associated projection matrices

$$P_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \widetilde{F}_m & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad P_6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \widetilde{S}_m \end{pmatrix}, \tag{3.25}$$

and with multiplicities given by (3.24). Upon introducing

$$S = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & v_m \\ 0 & 0 \end{pmatrix}, \qquad U = \begin{pmatrix} 0 & 0 \\ I_m & 0 \\ 0 & 0 \\ 0 & G_m \end{pmatrix}, \tag{3.26}$$

$$T = \begin{pmatrix} 1 & \lambda_2 \\ \lambda_2 & \frac{m+2}{m}\lambda_4 \end{pmatrix}, \qquad V = \begin{pmatrix} \lambda_2 & \lambda_4 \\ \lambda_4 & \frac{m+4}{m+2}\lambda_6 \end{pmatrix} \otimes I_m,$$

we obtain the representation

$$M = 2\lambda_4 P_4 + STS' + 6\lambda_6 P_6 + UVU'. \tag{3.27}$$

It is now easy to discuss the rank and the Moore–Penrose inverse of M.

Theorem 3.3:   The rank of a third order rotatable moment matrix M from (3.5) is

$$\text{rank } M = \begin{cases} 1 & \text{for } \lambda_2 = 0; \\ \binom{m+3}{3} - m - 1 & \text{for } \lambda_4 = \frac{m}{m+2}\lambda_2^2 > 0; \\ \binom{m+3}{3} - m & \text{for } \sqrt{\frac{m+4}{m+2}\lambda_2\lambda_6} = \lambda_4 > \frac{m}{m+2}\lambda_2^2 > 0; \\ \binom{m+3}{3} & \text{for } \sqrt{\frac{m+4}{m+2}\lambda_2\lambda_6} > \lambda_4 > \frac{m}{m+2}\lambda_2^2 > 0. \end{cases} \tag{3.28}$$

The rank is maximal, $\binom{m+3}{3}$, if and only if the moments satisfy

$$\frac{m+4}{m+2}\lambda_2\lambda_6 > \lambda_4^2; \tag{3.29}$$

and in this case the Moore–Penrose inverse is

$$M^+ = \frac{1}{2\lambda_4}P_4 + S(S'S)^{-1}T^{-1}(S'S)^{-1}S' + \frac{1}{6\lambda_6}P_6 + U(U'U)^{-1}V^{-1}(U'U)^{-1}U', \tag{3.30}$$

where P_4, P_6, S, T, U, V are given by (3.25+26).

Proof: Except for eigenvalues that vanish, the eigenvalues of STS' and of $S'ST$ are the same, as are those of UVU' and $U'UV$. The determinants are

$$\det S'ST = \det\begin{pmatrix} 1 & \lambda_2 \\ m\lambda_2 & (m+2)\lambda_4 \end{pmatrix}$$

$$= (m+2)\left(\lambda_4 - \frac{m}{m+2}\lambda_2^2\right), \tag{3.31}$$

$$\det U'UV = \det\begin{pmatrix} \lambda_2 & \lambda_4 \\ 3(m+2)\lambda_4 & 3(m+4)\lambda_6 \end{pmatrix} \otimes I_m$$

$$= 3^m\left((m+4)\lambda_2\lambda_6 - (m+2)\lambda_4^2\right)^m. \tag{3.32}$$

This implies (3.28). The formula (3.30) is verified directly using (3.27).   □

The four cases for the rank in (3.28) correspond to the number of spheres that the design $\tau_{\alpha,r,R}$ from (3.13) uses to realize M, namely 1/2 when $0 = r = R$, or 1 when $0 < r = R$, or 3/2 when $0 = r < R$, or 2 when $0 < r < R$.

A moment matrix M is often depicted by its associated information surface i_M, defined by

$$t \mapsto i_M(t) = \left(f(t)'M^+f(t)\right)^{-1} \tag{3.33}$$

for $f(t) \in \text{range } M$, and 0 otherwise. If the moment matrix is rotatable then the information surface becomes a function of the squared norm $t't$. Our representation keeps the two portions that arise from the two interlacing blocks in (3.5) separate.

Theorem 3.4:   If the moments satisfy (3.29), then the rotatable moment matrix M of (3.5) induces the rotatable information surface which depends on $t \in \mathbb{R}^m$ through $\rho^2 = t't$ according to

$$i_M(t) = \left(\frac{(m+2)\lambda_4 - 2\lambda_2\rho^2 + \frac{1}{2}[(m+1) - (m-1)\lambda_2^2/\lambda_4]\rho^4}{(m+2)\lambda_4 - m\lambda_2^2} \right. \tag{3.34}$$
$$\left. + \frac{(m+4)\lambda_6\rho^2 - 2\lambda_4\rho^4 + \frac{1}{6}[(m+1)\lambda_2 - (m-1)\lambda_4^2/\lambda_6]\rho^6}{(m+4)\lambda_2\lambda_6 - (m+2)\lambda_4^2} \right)^{-1}.$$

Proof: The rank of M is maximal, by Theorem 3.3, whence $i_M(t)$ is given by (3.33). Straightforward but lengthy evaluation of (3.30) yields (3.34).   □

The information surface (3.34) does *not* depend on the specific coordinate representation chosen to evaluate it. We find it much easier to calculate it via the Kronecker representation (1.2). For the representation in terms of two spherical uniform distributions, (3.34) reduces to (2.12) of Galil and Kiefer (1979).

We now discuss under what circumstances designs with rotatable moment matrices can be considered optimal.

4   Design Optimality

A broad optimality concept, not tied to a single scalar criterion, is the Kiefer ordering $\gg$ of Pukelsheim (1993, page 354). This is a two stage ordering, capturing an increase in symmetry relative to the congruence action of the underlying group $\mathcal{Q}$, as well as an improvement in the usual Loewner ordering of symmetric matrices.

Given an arbitrary third order moment matrix A, an increase in symmetry is achieved by averaging it over the orbit under the action of the group $\mathcal{Q}$,

$$\overline{A} = \int_{\mathrm{Orth}(m)} Q_R A Q'_R \, dR.$$

The Haar probability measure dR on the orthogonal group $\mathrm{Orth}(m)$ is not easy to handle. However, the average $\overline{A}$ coincides with the projection of A, onto the 4-dimensional subspace of symmetric matrices that have the pattern (3.5); see Pukelsheim (1993, page 349). This opens up a way of calculating $\overline{A}$.

If the projection $\overline{A}$ is represented as in (3.5) and has coefficient $\lambda_2 \in (0,1)$, a maximization of λ_6 within its range (3.4c) gives a Loewner improvement which is best for given values λ_2 and λ_4. Attainment of the upper bound

$$\lambda_6^M = \frac{m+2}{m+4} \, \frac{(\lambda_4 - \frac{m}{m+2}\lambda_2^2) + \frac{m+2}{m}\lambda_4(\frac{m}{m+2}\lambda_2 - \lambda_4)}{1 - \lambda_2} \tag{4.1}$$

forces attainment of the upper bound in (3.12c), and a squared radius $R^2 = 1$ in (3.17). Accordingly we set $R = 1$ in (3.13–17), and introduce the two-parameter family of *boundary nucleus designs*

$$\tau_{\alpha,r} = \alpha u_{\sqrt{m}} + (1-\alpha)u_{r\sqrt{m}}, \qquad \alpha \in [0,1], \ r \in [0,1]. \tag{4.2}$$

They allocate weight α on the boundary sphere of the experimental domain $\mathcal{T}_{\sqrt{m}}$, while placing the remaining weight $1-\alpha$ on a concentric inner sphere nucleus of radius $r\sqrt{m}$.

The different cases for the rank distinction (3.28) thus correspond to $\alpha = 0$ and $r = 0$ ($\frac{1}{2}$ sphere), or $\alpha = 0$ and $r > 0$ (1 sphere), or $\alpha \in (0,1)$ and $r = 0$ ($\frac{3}{2}$ spheres), or $\alpha \in (0,1)$ and $r \in (0,1)$ (2 spheres).

The boundary nucleus designs form a complete class in the Kiefer ordering, as follows.

Theorem 4.1: *Given any design τ on the experimental domain $\mathcal{T}_{\sqrt{m}}$, there always exist some boundary weight $\alpha \in [0,1]$ and some nucleus radius $r \in [0,1]$ such that the boundary nucleus design $\tau_{\alpha,r}$ improves upon τ in the Kiefer ordering, $M(\tau_{\alpha,r}) \gg M(\tau)$.*

Proof: For the 4-dimensional space of symmetric matrices which contains the rotatable moment matrices (3.5), we use the orthogonal basis

$$
W_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad
W_2 = \begin{pmatrix} 0 & 0 & v'_m & 0 \\ 0 & I_m & 0 & 0 \\ v_m & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
$$
$$
W_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & G'_m \\ 0 & 0 & F_m & 0 \\ 0 & G_m & 0 & 0 \end{pmatrix}, \qquad
W_6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & S_m \end{pmatrix}.
$$
$$(4.3)$$

Let A be the moment matrix of τ. We calculate its projection onto the symmetric matrices of pattern (3.5) through $\overline{A} = \sum_{i=0,2,4,6} \lambda_i W_i$, with coefficients given by

$$\lambda_i = \frac{\text{trace } AW_i}{\text{trace } W_i^2}. \tag{4.4}$$

We have trace $W_0^2 = 1$, trace $W_2^2 = 3m$, trace $W_4^2 = 9m(m+2)$, trace $W_6^2 = 15m(m+2)(m+4)$.

In (4.4) we always get $\lambda_0 = 1$. In case $\lambda_4 = \frac{m}{m+2}\lambda_2^2$, the matrix $\overline{A}$ is the moment matrix of the degenerate boundary nucleus design $\tau_{0,r}$ with radius $r = \sqrt{\lambda_2} \in [0,1]$.

The case $\lambda_4 > \frac{m}{m+2}\lambda_2^2$ necessitates $\lambda_2 \in (0,1)$. With λ_6^M from (4.1), we introduce $M = W_0 + \lambda_2 W_2 + \lambda_4 W_4 + \lambda_6^M W_6$. This is a Loewner improvement over $\overline{A}$, $M \geq \overline{A}$. Furthermore M is attained by a boundary nucleus design $\tau_{\alpha,r}$, as shown in the proof of Theorem 3.1. The nucleus radius r is found using a, b, c from (3.10) which, in the present case, satisfy $b - a^2 > 0$. Hence in (3.15), the first equation entails $\alpha \neq 0, 1$ and $r^2 \neq 1$, while the last equation yields $r^2 = (ac - b^2)/(b - a^2)$. Together with the boundary weight α from (3.17) we obtain

$$r^2 = \frac{m+2}{m} \frac{\frac{m+4}{m+2}\lambda_2 \lambda_6^M - \lambda_4^2}{\lambda_4 - \frac{m}{m+2}\lambda_2^2} \in [0,1), \qquad \alpha = \frac{\lambda_2 - r^2}{1 - r^2} \in (0,1). \qquad \square$$

It is an immediate consequence that the boundary nucleus designs contain an overall ϕ-optimal design, under every scalar criterion ϕ which is isotonic relative to the Kiefer ordering. A sufficient set of conditions is the following.

Corollary 4.2:   Let the criterion $\phi : \mathrm{NND}(k) \mapsto \mathbb{R}$ be Loewner isotonic, concave, and $\mathcal{Q}$-invariant. Then there is a boundary nucleus design $\tau_{\alpha,r}$ that is ϕ-optimal, in the set of all designs on the experimental domain $\mathcal{T}_{\sqrt{m}}$. If, in addition, the criterion ϕ is strictly concave then every ϕ-optimal design must be a boundary nucleus design.   □

$\mathcal{Q}$-invariance means $\phi(M) = \phi(Q_R M Q_R')$, for all matrices Q_R in (1.3). Since $\mathcal{Q}$ is a subgroup of the orthogonal group $\mathrm{Orth}(k)$, orthogonal invariance implies $\mathcal{Q}$-invariance. An orthogonally invariant criterion depends on M only through the eigenvalues of M. The most important criteria of this sort are the matrix means ϕ_p with $p \in [-\infty, 1]$, see Pukelsheim (1993, page 140). With $p = 0, -1, -\infty$ they comprise the classical D-, A-, and E-criteria. For rotatable moment matrices M, we have studied the eigenvalues in the previous section. We shall use these results to investigate ϕ_p-optimal designs more closely, in a companion paper.

Here we use Corollary 4.2 to show that the boundary nucleus designs generate precisely those moment matrices that are admissible in the set of *all* moment matrices. The admissibility discussion for the usual minimal set of regression functions is more complicated, see Section 2 of Heiligers and Schneider (1992). A moment matrix M is called admissible when every competing moment matrix A with $A \geq M$ is actually equal to M, see Pukelsheim (1993, page 247).

Theorem 4.3:   In the third order model with regression function f given by (1.2), the moment matrix of a boundary nucleus design is admissible.

Proof: The moment matrix M of a boundary nucleus design has pattern (3.5), with $\lambda_6 = \lambda_6^M$ from (4.1). Let A be an arbitrary moment matrix satisfying $A \geq M$. From $Q_R A Q_R' \geq Q_R M Q_R' = M$, for all $R \in \mathrm{Orth}(m)$, we see that the projection $\overline{A} = \int Q_R A Q_R' \, dR$ of A onto the rotatable moment matrices fulfills $\overline{A} \geq M$. Theorem 3.1 provides scalars μ_2, μ_4, μ_6 such that

$$\overline{A} = \begin{pmatrix} 1 & 0 & \mu_2 v_m' & 0 \\ 0 & \mu_2 I_m & 0 & \mu_4 G_m' \\ \mu_2 v_m & 0 & \mu_4 F_m & 0 \\ 0 & \mu_4 G_m & 0 & \mu_6 S_m \end{pmatrix}. \tag{4.5}$$

Premultiplication of $\overline{A} - M$ by $(1, 0, \delta v_m', 0)$ and postmultiplication by the transposed vector gives

$$0 \leq 2\delta(\mu_2 - \lambda_2)m + \delta^2(\mu_4 - \lambda_4)m(m+2), \qquad \text{for all } \delta \in \mathbb{R}.$$

This necessitates $\mu_2 = \lambda_2$. Next, premultiplication of $\overline{A} - M$ by $(0, I_m, 0, \delta G'_m)$ and postmultiplication by its transpose yields

$$0 \le \Big(\delta(\mu_4 - \lambda_4) + \delta^2(\mu_6 - \lambda_6)\big[6 + 3(m+2)\big]\Big) 3(m+2)\, I_m, \qquad \text{for all } \delta \in \mathbb{R}.$$

This forces $\mu_4 = \lambda_4$. Finally $\mu_6 \ge \lambda_6 = \lambda_6^M$ entails $\mu_6 = \lambda_6$. This proves $M = \overline{A}$. Together with $A \ge M$ we see that the difference $A - \overline{A}$ is nonnegative definite. On the other hand we have $\text{trace}(A - \overline{A}) = \text{trace}\, A - \int_{\text{Orth}(m)} \text{trace}\, Q'_R Q_R A \, dR = 0$. Now $A = \overline{A} = M$, which shows that M is admissible. $\qquad \square$

Conversely, a rotatable moment matrix that is admissible among all rotatable moment matrices is attained by a boundary nucleus design. These admissibility results and the arguments to establish them parallel those of Lemma 1 of Heiligers (1991) and Lemma 2 of Heiligers and Schneider (1992), for the case of a minimal system of components.

5   Third Order Rotatability of Experimental Designs

As with second order rotatable designs, third order rotatable designs can always be formed by combining symmetric sets of design points on concentric spheres. Many such combinations are feasible, and the symmetric sets themselves can be arrived at in various ways. It is often hard to select a design with a suitably small number of points due to the several moment conditions that must be satisfied simultaneously. Specific discussions are given by Adhikary and Panda (1984, 1985, 1986), Arap Koske and Patel (1989), Bagchi (1986), Draper (1960a,b; 1961), Gardiner, Grandage and Hader (1959), Herzberg (1964, 1967), Huda (1981, 1982a,b; 1983, 1984, 1985, 1987a,b; 1988, 1989, 1991), Mukerjee and Huda (1985), Narasimham and Rao (1980), Nigam (1967), Panda (1982), Panda and Das Roy (1990b), Thaker and Das (1961), and Tyagi (1964).

Published practical applications are harder to find. It is unusual in practical work to fit a third order surface. More often, the response data are transformed in order to avoid such a fit, and to reduce the fitted model to second order. Practical applications are described by Baker and Bargmann (1985), and Derringer (1969).

6   Higher Order Rotatability

Theoretical interest in higher order rotatability has led to a number of papers. For fourth order rotatability, see Arap Koske (1987, 1989), Arap Koske and Patel (1986, 1987, 1989), Draper and Herzberg (1985), Huda (1991), Huda and Mukerjee (1989), Huda and Shafiq (1987), Mukerjee (1987), Mukerjee and Huda (1990), Panda and Das Roy (1990a), Patel and Arap Koske (1985), Shafiq and Huda (1989). For fifth order rotatability, see Njui and Patel (1988).

For the extension of our results to rotatability of arbitrary order, a key result is part (iii) of Theorem 3.1, in that it extracts the pure matrix algebra aspects of the problem. This result has close relationships to tensor product representations of the orthogonal group, see Brauer (1937), Wales (1987). It is instructive to contemplate these interrelations in some greater detail.

The group representation $R \mapsto R^{\otimes d}$ is called the dth order tensor representation, and takes values in the vectorspace $\mathbb{R}^{m^d \times m^d}$ of real $m^d \times m^d$ matrices. Its commutating algebra is defined by

$$
\begin{aligned}
\mathcal{B}_d^{(m)} &= \{B \in \mathbb{R}^{m^d \times m^d} : R^{\otimes d} B = B R^{\otimes d} \quad \text{for all } R \in \text{Orth}(m)\} \\
&= \{B \in \mathbb{R}^{m^d \times m^d} : R^{\otimes d} B R^{\otimes d\prime} = B \quad \text{for all } R \in \text{Orth}(m)\},
\end{aligned}
\tag{6.1}
$$

assembling those square matrices B which commute with $R^{\otimes d}$, that is, which are invariant under the action by conjugacy. Brauer (1937) was the first to investigate this algebra more closely, and he proved that its dimension b_d is given by the odd factorials,

$$
\begin{aligned}
b_d &= (2d - 1)\, b_{d-1} \\
&= (2d - 1) \cdot (2d - 3) \cdot \ldots \cdot 5 \cdot 3 \cdot 1.
\end{aligned}
\tag{6.2}
$$

This follows from a familiar combinatorial argument: The terms in the sum of (2.4) generally take the form

$$
\sum_{i_1,\ldots,i_d,j_1,\ldots,j_d} e_{i_1,\ldots,i_d} e'_{j_1,\ldots,j_d},
\tag{6.3}
$$

subject to the restriction that the $2d$ subscripts $i_1, \ldots, i_d, j_1, \ldots, j_d$ form what is called a *matching*, that is, each of the numbers $1, \ldots, d$ appears exactly twice, see Section 5 of Brauer (1937). Hence given any number, say d, there are $2d - 1$ positions for its partner, times b_{d-1} matchings of the remaining numbers $1, \ldots, d - 1$. This proves (6.2). For $d = 0, 1, 2, 3, 4$ we get $b_d = 1, 1, 3, 15, 105$.

In our statistical problem all matrices are symmetric, whence the nonsymmetric matrices in $\mathcal{B}_d^{(m)}$ are redundant. We intersect $\mathcal{B}_d^{(m)}$ with the space of symmetric

matrices, and define

$$\mathcal{S}_d^{(m)} = \{S \in \mathbb{R}^{m^d \times m^d} : S = S', \quad R^{\otimes d}S = SR^{\otimes d} \quad \text{for all } R \in \text{Orth}(m)\}$$
$$= \{S \in \mathbb{R}^{m^d \times m^d} : S = S', \quad R^{\otimes d}SR^{\otimes d\prime} = S \quad \text{for all } R \in \text{Orth}(m)\}. \tag{6.4}$$

We claim that its dimension s_d obeys the recurrence relation

$$s_d = s_{d-1} + (2d-2)s_{d-2} + \frac{b_d - s_{d-1} - (2d-2)s_{d-2}}{2}, \tag{6.5}$$

with starting values $s_0 = s_1 = 1$. To see this, we call a matching *symmetric* when the associated matrix in (6.3) is symmetric. From the set of all matchings we extract the symmetric ones, in two steps. First we pair off i_1 and j_1, leaving s_{d-1} ways to complete the whole matching symmetrically. Second we pair off i_1 with any one of the $2d-2$ symbols $i_2, \ldots, i_d, j_2, \ldots, j_d$; this determines a countercrosswise pairing for j_1 to preserve symmetry, and then leaves s_{d-2} possible completions. Thirdly, there remain $b_d - s_{d-1} - (2d-2)s_{d-2}$ matchings for which the matrix (6.3) is *not* symmetric. Each of these must come together with its "transposed" counterpart, so that the two associated matrices become symmetric as a sum. (For instance, in (2.4) the four double sums originate in this way.) This proves (6.5). For $d = 0, 1, 2, 3, 4$ we have $s_d = 1, 1, 3, 11, 67$.

The emphasis on the symmetric commutator algebra $\mathcal{S}_d^{(m)}$ is one point where the statistical problem appears to deviate from the group representation literature. A second point is that the rotatable symmetric matrices in (2.1) also involve all lower order representations. That is, we actually deal with the Cartesian product of *all* tensor representations up to order d,

$$R \mapsto (1, R, R^{\otimes 2}, \ldots, R^{\otimes d}).$$

This has subtle consequences. To fix ideas we consider $d = 4$. The matrix corresponding to (2.1) then involves an $m^4 \times m^2$ off-diagonal block $T(\theta)$, say. When this block is rearranged as a square $m^3 \times m^3$ matrix, it need *not* be symmetric. Hence $T(\theta)$ contributes, not $s_3 = 11$, but $b_3 = 15$ degrees of freedom. The following matrix displays the number of parameters that are called for by each block of a fourth order rotatable symmetric matrix,

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 3 \\ & 1 & 0 & 3 & 0 \\ & & 3 & 0 & 15 \\ & & & 11 & 0 \\ & & & & 67 \end{pmatrix}. \tag{6.6}$$

More generally, let p_d be the dimension of the subspace of d th order rotatable matrices, in the space of symmetric $(1+m+m^2+m^3+m^4) \times (1+m+m^2+m^3+m^4)$ matrices. For $d = 0, 1, 2, 3, 4$ we then have $p_d = 1, 2, 6, 20, 105$.

The rapid increase in dimensionality is much dampened when it comes to rotatability of moment matrices. Even the Moore–Penrose inverses do not exhaust all p_d degrees of freedom. For instance, the third order rotatable Moore–Penrose inverse of (3.30) can be represented using, not $p_3 = 20$, but just 8 scalars $\alpha, \beta, \gamma, \delta_1, \delta_2, \eta, \epsilon_1, \epsilon_2$,

$$
\begin{pmatrix}
\alpha & 0 & \gamma v'_m & 0 \\
0 & \beta I_m & 0 & \eta G'_m \\
\gamma v_m & 0 & \delta_1 \Pi_4 + \delta_2 v_m v'_m & 0 \\
0 & \eta G_m & 0 & \epsilon_1 \Pi_6 + \epsilon_2 G_m G'_m
\end{pmatrix}. \tag{6.7}
$$

The reason is that the matrices with this pattern form a quadratic subspace of symmetric matrices, compare page 132 of DGP, or Lemma 13.10 of Pukelsheim (1993). The dimension of this subspace is reduced by symmetrizing the d-fold Kronecker products as in (3.21). Let q_d be the dimension of the *smallest* quadratic subspace of symmetric matrices that contains the d th order rotatable moment matrices. For $d = 0, 1, 2, 3$ our results show that $q_d = 1, 2, 5, 8$.

References

Adhikary B, Panda R (1984) Group divisible third order rotatable designs (GDTORD). Sankhyā B46:135–146

Adhikary B, Panda R (1985) Group divisible response surface (GDRS) designs of third order. Calcutta Statist Assoc Bull 34:75–87

Adhikary B, Panda R (1986) Construction of nearly D-efficient third order rotatable designs using PBIB designs. Symposium on Optimization, Design of Experiments and Graph Theory, IIT, Bombay

Arap Koske JK (1987) A fourth order rotatable design in four dimensions. Comm Statist A16:2747–2753

Arap Koske JK (1989) The variance function of the difference between two estimated fourth order response surface. J Statist Plann Inference 23:263–266

Arap Koske JK, Patel MS (1986) A fourth order rotatable design in three dimensions. Comm Statist A15:3435–3444

Arap Koske JK, Patel MS (1987) Construction of fourth order rotatable designs with estimation of corresponding response surface. Comm Statist A16:1361–1376

Arap Koske JK, Patel MS (1989) A simpler way of obtaining non-singularity conditions of rotatability. Comm Statist A18:2489–2500

Bagchi S (1986) A series of nearly D-optimal third order rotatable designs. Sankhyā B48:186–198

Baker FD, Bargmann RE (1985) Orthogonal central composite designs of the third order in the evaluation of sensitivity and plant growth simulation models. J Amer Statist Assoc 80:574–579

Box GEP, Hunter JS (1957) Multi-factor experimental designs for exploring response surfaces. Ann Math Statist 28:195–241

Brauer R (1937) On algebras which are connected with the semisimple continuous groups. Ann of Math 38:857–872

Derringer GC (1969) Sequential method for estimating response surfaces. Indus Eng Chem 61:6–13

Draper NR (1960a) Third order rotatable designs in three dimensions. Ann Math Statist 31:865–874

Draper NR (1960b) A third order rotatable design in four dimensions. Ann Math Statist 31:875–877

Draper NR (1961) Third order rotatable designs in three dimensions: some specific designs. Ann Math Statist 32:910–913

Draper NR (1984) Schläflian rotatability. J Roy Statist Soc B46:406–411

Draper NR, Herzberg AM (1985) Fourth order rotatability. Comm Statist B14:515–528

Draper NR, Pukelsheim F (1990) Another look at rotatability. Technometrics 32:195–202

Draper NR, Gaffke N, Pukelsheim F (1991) First and second order rotatability of experimental designs, moment matrices, and information surfaces. Metrika 38:129–161

Farrell RH, Kiefer J, Walbran A (1967) Optimum multivariate designs. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1:113–138 (Also in: Jack Carl Kiefer Collected Papers III: Design of Experiments. Springer, New York 1985, pp 247–272)

Gaffke N, Heiligers B (1992) Computing optimal approximate invariant designs for cubic regression on multidimensional balls and cubes. Report No. 365, Institut für Mathematik, Universität Augsburg.

Galil Z, Kiefer JC (1977) Comparison of rotatable designs for regression on balls I (quadratic). J Statist Plann Inference 1:27–40 (Also in: Jack Carl Kiefer Collected Papers III: Design of Experiments. Springer, New York 1985, pp 391–404)

Galil Z, Kiefer JC (1979) Extrapolation designs and Φ_p-optimum designs for autoregression on the q-ball. J Statist Plann Inference 3:27–38 (Also in: Jack Carl Kiefer Collected Papers III: Design of Experiments. Springer, New York 1985, pp 467–478)

Gardiner DA, Grandage AHE, Hader RJ (1959) Third order rotatable designs for exploring response surfaces. Ann Math Statist 30:1082–1096

Heiligers B (1991) Admissibility of experimental designs in linear regression with constant term. J Statist Plann Inference 28:107–123

Heiligers B, Schneider K (1992) Invariant admissible and optimal designs in cubic regression on the ν-ball. J Statist Plann Inference 31:113–125

Herzberg AM (1964) Two third order rotatable designs in four dimensions. Ann Math Statist 35:445–446

Herzberg AM (1967) Cylindrically rotatable designs of types 1, 2, and 3. Ann Math Statist 38:167–176

Huda S (1981) Cylindrically rotatable designs of type 3: further considerations. Biometrical J 24:469–475

Huda S (1982a) Some third order rotatable designs in three dimensions. Ann Inst Statist Math 34:365–371

Huda S (1982b) Some third order rotatable designs. Biometrical J 24:257–263

Huda S (1983) Two third-order rotatable designs in four dimensions. J Statist Plann Inference 8:241–243

Huda S (1984) On D-efficiency of some third-order rotatable designs. J Indian Soc Agricultural Statist 36:51–67

Huda S (1985) Some 212-point third-order rotatable designs in six dimensions. J Statist Res 19:63–64

Huda S (1987a) The construction of third-order rotatable designs in k dimensions from those in lower dimensions. Pakistan J Statist 3A:11–16

Huda S (1987b) Mixed-cylindrically rotatable designs. Pakistan J Statist 3A:63–67

Huda S (1988) A note on the analysis of third-order cylindrically rotatable designs of type 3. Pakistan J Statist 4A:139–146

Huda S (1989) The m-grouped cylindrically rotatable designs of types $(1, 0, m - 1), (0, 1, m - 1), (1, 1, m - 2)$ and $(0, 0, m)$. Pakistan J Statist 5A:109–117

Huda S (1991) On some D_s-optimal designs in spherical regions. Comm Statist A20:2965–2985

Huda S, Mukerjee R (1989) D-optimal measures for fourth-order rotatable designs. Statistics 20:353–356

Huda S, Shafiq M (1987) On D_s-efficiency of D-optimal fourth-order rotatable designs. Pakistan J Statist 3B:33–37

Karlin S, Studden WJ (1966) Tchebycheff Systems: With Applications in Analysis and Statistics. Wiley-Interscience, New York

Khuri AI (1988) A measure of rotatability for response-surface designs. Technometrics 30:95–104

Khuri AI (1992) Diagnostic results concerning a measure of rotatability. J Roy Statist Soc B54:253–267

Kiefer JC (1960) Optimum experimental designs V with applications to systematic and rotatable designs. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability 1:381–405 (Also in: Jack Carl Kiefer Collected Papers III: Design of Experiments. Springer, New York 1985, pp 103–127)

Mukerjee R (1987) On fourth-order rotatable designs. Comm Statist A16:1697–1702

Mukerjee R, Huda S (1985) Minimax second- and third-order designs to estimate the slope of a response surface. Biometrika 72:173–178

Mukerjee R, Huda S (1990) Fourth-order rotatable designs: A-optimal measures. Statist Prob Lett 10:111–117

Narasimham VL, Rao KN (1980) A modified method for the construction of third order rotatable designs through a pair of balanced incomplete block designs. Proc Second Annual Conf ISTPA, Bombay December 1980.

Nigam AK (1967) On third order rotatable designs with smaller number of levels. J Indian Soc Agricultural Statist 19:36–41

Njui F, Patel MS (1988) Fifth order rotatability. Comm Statist A17:833–848

Panda R (1982) Contributions to Response Surface Designs. PhD Thesis, Calcutta University

Panda R, Das Roy A (1990a) Analysis of fourth order rotatability in k-dimensions. Calcutta Statist Assoc Bull 39:195–200

Panda R, Das Roy A (1990b) Group divisible third order rotatable designs in non-orthogonal blocks. J Indian Soc Agricultural Statist 42:189–200

Patel MS, Arap Koske JK (1985) Conditions for fourth order rotatability in k dimensions. Comm Statist A14:1343–1351

Pukelsheim F (1980) Multilinear estimation of skewness and kurtosis in linear models. Metrika 27:103–113

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

Shafiq M, Huda S (1989) On application of association matrices in the analysis of fourth-order rotatable designs. Pakistan J Statist 5A:131–142

Thaker PJ, Das MN (1961) Sequential third order rotatable designs for up to eleven factors. J Indian Soc Agricultural Statist 13:218–231

Tyagi BN (1964) On construction of second and third order rotatable designs through pair-wise balanced and doubly balanced designs. Calcutta Statist Assoc Bull 13:150–162

Wales D (1987) Eigenvalues connected to the radical of Brauer's centralizer algebras. The Arcata Conference on Representations of Finite Groups (P Fong, Ed). Proceedings of Symposia in Pure Mathematics 47,2:547–552