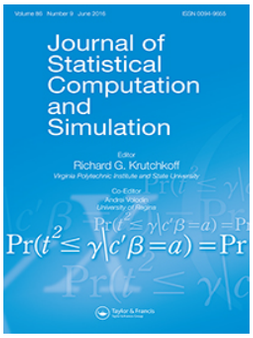

Journal of Statistical Computation and Simulation

ISSN: 0094-9655 (Print) 1563-5163 (Online) Journal homepage: http://www.tandfonline.com/loi/gscs20

Enhanced design efficiency through least upper bounds

D.R. Jensen & D.E. Ramirez

To cite this article: D.R. Jensen & D.E. Ramirez (2016) Enhanced design efficiency through least upper bounds, Journal of Statistical Computation and Simulation, 86:9, 1798-1817, DOI: 10.1080/00949655.2015.1082133

To link to this article: http://dx.doi.org/10.1080/00949655.2015.1082133

Journal of Statistical Computation and Simulation, 2016
Vol. 86, No. 9, 1798–1817, http://dx.doi.org/10.1080/00949655.2015.1082133

Enhanced design efficiency through least upper bounds

D.R. Jensen[a] and D.E. Ramirez[b]*

[a]*Department of Statistics, Virginia Tech, Blacksburg, VA, USA;* [b]*Department of Mathematics, University of Virginia, Charlottesville, VA, USA*

(*Received 25 February 2015; accepted 9 August 2015*)

Lower and upper *spectral* bounds are known for positive-definite $(k \times k)$ matrices in $(\mathbb{S}_k^+, \succeq_L)$ under Loewner (Uber monotone Matrixfunktionen. Math Z. 1934;38:177–216) ordering. Lower and upper *singular* bounds for matrices of order $(n \times k)$ in $(\mathbb{F}_{n\times k}, \succeq)$ derive under an induced ordering. These orderings are combined here to the following effects. Given two first-order experimental designs $(\boldsymbol{X}, \boldsymbol{Z})$ in $(\mathbb{F}_{n\times k}, \succeq)$, their *upper singular bound* $\boldsymbol{X}_M$ enhances both $\boldsymbol{X}$ and $\boldsymbol{Z}$ in that its Fisher Information matrix dominates those for both $\boldsymbol{X}$ and $\boldsymbol{Z}$, thus ordering essentials in Gauss–Markov estimation. Moreover, if $\boldsymbol{\Sigma}$, $\boldsymbol{\Omega}$, and $\boldsymbol{\Xi}$ are dispersion matrices for linear estimators under $\boldsymbol{X}, \boldsymbol{Z}$, and $\boldsymbol{X}_M$, respectively, then $\boldsymbol{\Xi}$ is the *spectral lower bound* for $(\boldsymbol{\Sigma}, \boldsymbol{\Omega})$ in $(\mathbb{S}_k^+, \succeq_L)$. In essence this algorithm identifies elements in $\boldsymbol{Z}$ complementary to those of $\boldsymbol{X}$, and combines these into $\boldsymbol{X}_M$. Case studies illustrate gains to be made thereby in first and second-order designs. Specifically, two examples demonstrate that designs optimal under separate criteria may be combined into a single design dominating both. In addition, selected examples demonstrate that classical second-order designs may be improved *inter se*.

Keywords: symmetrical and rectangular matrix extremes; monotone functions; efficiency indices; design augmentation

1. Introduction

Extremal problems have pervaded applied probability and statistics from their beginnings. These include maximal, minimal, and optimal solutions as foundation stones. Such solutions often shed new light on structural aspects of the systems at hand.

Specifically, *spectral* lower $(\boldsymbol{A} \wedge \boldsymbol{B})$ and upper $(\boldsymbol{A} \vee \boldsymbol{B})$ bounds are known for positive-definite matrices $(\boldsymbol{A}, \boldsymbol{B})$ in $(\mathbb{S}_k^+, \succeq_L)$ under Loewner [1] ordering. In addition, lower $(\boldsymbol{X} \wedge \boldsymbol{Z})$ and upper $(\boldsymbol{X} \vee \boldsymbol{Z})$ *singular* bounds derive for matrices $(\boldsymbol{X}, \boldsymbol{Z})$ in $(\mathbb{F}_{n\times k}, \succeq)$ under an invariant ordering induced on $\mathbb{F}_{n\times k}$. Details are found in [2]. These orderings are combined here to the following effects. Given two first-order experimental designs $(\boldsymbol{X}, \boldsymbol{Z})$ in $(\mathbb{F}_{n\times k}, \succeq)$, their *upper singular bound* $\boldsymbol{X}_M$ enhances both $\boldsymbol{X}$ and $\boldsymbol{Z}$ in that its Fisher Information matrix dominates those for both $\boldsymbol{X}$ and $\boldsymbol{Z}$, thus ordering essentials in Gauss–Markov procedures. These gains are achieved formally on isolating elements of $\boldsymbol{Z}$ that serve to complement those of $\boldsymbol{X}$, and combining those elements with $\boldsymbol{X}$ to obtain $\boldsymbol{X}_M$. Moreover, if $\boldsymbol{\Sigma}, \boldsymbol{\Omega}$, and $\boldsymbol{\Xi}$ are dispersion matrices for linear estimators under $\boldsymbol{X}, \boldsymbol{Z}$, and $\boldsymbol{X}_M$, respectively, then $\boldsymbol{\Xi}$ is the *spectral lower bound* for $(\boldsymbol{\Sigma}, \boldsymbol{\Omega})$ in $(\mathbb{S}_k^+, \succeq_L)$, i.e. $\boldsymbol{\Xi} = \min(\boldsymbol{\Sigma}, \boldsymbol{\Omega})$.

*Corresponding author. Email: der@virginia.edu

In a standard design using $n = 8$ runs, a subsequent example demonstrates that half the design, when modified as prescribed here, yields estimates for linear and second-order coefficients more precise than the original design with eight runs. An outline of the paper follows.

Preliminary developments in Section 2 establish notation and essentials for orderings on $(\mathbb{S}_k^+, \succeq_L)$ and $(\mathbb{F}_{n\times k}, \succeq)$. Our main results are developed in Section 3, where two first-order experimental designs $(\boldsymbol{X}, \boldsymbol{Z})$ combine to give their *upper singular bound* $\boldsymbol{X}_M$, its Fisher Information matrix dominating those for both $\boldsymbol{X}$ and $\boldsymbol{Z}$. Essentials in Gauss–Markov procedures are thereby ordered as documented in detail. Applications in Section 4 serve to illustrate these findings through selected case studies.

2. Preliminaries

Conventions for notation are followed by basic properties of $(\mathbb{S}_k^+, \succeq_L)$ and $(\mathbb{F}_{n\times k}, \succeq)$, as well as functions monotone on these spaces. The principal features summarized here are covered in detail in [2]; two essential theorems from that work are stated without proof in the Appendix, as the cited *Proceedings* source may not be widely available.

2.1. *Notation*

Denote by $\mathbb{R}^k$ the Euclidean k-space and by $\mathbb{R}_+^k$ its positive orthant; by $\mathbb{F}_{n\times k}$ the real $(n \times k)$ matrices of rank $k \le n$; by $\mathbb{S}_k$ the real symmetric $(k \times k)$ matrices, with $\mathbb{S}_k^0$, $\mathbb{S}_k^+$ and $\mathbb{D}_k$ as their positive semidefinite, positive definite, and diagonal varieties. The transpose, trace, and determinant of $\boldsymbol{A}$ are denoted by $\boldsymbol{A}'$, $\mathrm{tr}(\boldsymbol{A})$, and $|\boldsymbol{A}|$ where applicable; and special arrays include the unit vector $\mathbf{1}_k = [1, \ldots, 1]' \in \mathbb{R}^k$, the unit matrix $\boldsymbol{I}_k$, and a typical diagonal matrix $\boldsymbol{D}_\alpha = \mathrm{Diag}(\alpha_1, \ldots, \alpha_k) \in \mathbb{D}_k$. Transformation groups acting on $\mathbb{R}^k$ include the general linear group $\mathcal{G}l(k)$ and the real orthogonal group $\mathcal{O}(k)$; and elements of $\mathbb{H}_{n\times k}$ comprise the semiorthogonal matrices $\boldsymbol{H}$ such that $\boldsymbol{H}\boldsymbol{H}'$ is idempotent of rank k and $\boldsymbol{H}'\boldsymbol{H} = \boldsymbol{I}_k$. The *spectral decomposition* $\boldsymbol{A} = \sum_{i=1}^k \alpha_i \boldsymbol{q}_i \boldsymbol{q}_i' \in \mathbb{S}_k^+$ yields its *symmetric root* $\boldsymbol{A}^{\frac{1}{2}} = \sum_{i=1}^k \alpha_i^{\frac{1}{2}} \boldsymbol{q}_i \boldsymbol{q}_i'$. The *singular decomposition* of $\boldsymbol{X} \in \mathbb{F}_{n\times k}$ is $\boldsymbol{X} = \sum_{i=1}^k \lambda_i \boldsymbol{p}_i \boldsymbol{q}_i' = \boldsymbol{P}\boldsymbol{D}_\lambda \boldsymbol{Q}'$ in which $\boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_k] \in \mathbb{H}_{n\times k}$ contains the *left singular vectors,* $\boldsymbol{Q} = [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_k] \in \mathcal{O}(k)$ contains the *right singular vectors,* and $\boldsymbol{D}_\lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_k)$, with $\{\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_k\}$, comprise the ordered *singular values* of $\boldsymbol{X}$.

Standard usage refers to independent, identically distributed *(iid)* variates, their cumulative distribution function *(cdf)* and $\mathcal{L}(\boldsymbol{Y})$ the distribution of $\boldsymbol{Y}$, with $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ as the Gaussian law on $\mathbb{R}^k$ having the mean $\mathrm{E}(\boldsymbol{Y}) = \boldsymbol{\mu}$ and dispersion matrix $\mathrm{V}(\boldsymbol{Y}) = \boldsymbol{\Sigma}$.

2.2. *Models*

Models $\{\boldsymbol{Y} = \beta_0 \mathbf{1}_n + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$ with intercept typically are encountered in practice, under assumptions to follow.

Assumptions

A1: $\mathrm{E}(\boldsymbol{\epsilon}) = \boldsymbol{0}$ *and* $\mathrm{V}(\boldsymbol{\epsilon}) = \sigma^2 \boldsymbol{I}_n$;
A2: $\mathcal{L}(\boldsymbol{\epsilon}) = N_n(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n)$.

The columns of $\boldsymbol{X}(n \times k)$ often are centered so that $\mathbf{1}_n'\boldsymbol{X} = \mathbf{0}$. Nonlinear models emerge on replacing $\boldsymbol{X}$ by some function $f(\boldsymbol{X})$, so that $\{\boldsymbol{Y} = \beta_0 \mathbf{1}_n + f(\boldsymbol{X})\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$. In what follows we distinguish carefully between the *design* matrix $\boldsymbol{X}$, and the *model matrix* $\mathbf{X}(\boldsymbol{X})$ as in the following.

Definition 1

(i) *Given a design* $\boldsymbol{X}$ *in* $\{\boldsymbol{Y}=\beta_0\boldsymbol{1}_n+f(\boldsymbol{X})\boldsymbol{\beta}+\boldsymbol{\epsilon}\}$, *the model matrix* $\mathbf{X}$ *incorporates both the design and model as* $\mathbf{X}=[\boldsymbol{1}_n,f(\boldsymbol{X})]$. *For another design* $\mathbf{Z}$ *in context, the model matrix is* $\mathbf{Z}=[\boldsymbol{1}_n,f(\mathbf{Z})]$. *retaining the functional form* $f(\cdot)$ *together with the new design.*

(ii) *The Fisher Information matrix for* $\boldsymbol{\beta}$ *is* $\sigma^2\mathcal{I}(\boldsymbol{\beta})=\mathbf{X}'\mathbf{X}$ *under design* $\boldsymbol{X}$, *and* $\sigma^2\mathcal{I}(\boldsymbol{\beta})=\mathbf{Z}'\mathbf{Z}$ *under design* $\mathbf{Z}$, *with inverses* $\mathrm{V}(\hat{\boldsymbol{\beta}})=\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ *and* $\mathrm{V}(\hat{\boldsymbol{\beta}})=\sigma^2(\mathbf{Z}'\mathbf{Z})^{-1}$ *under the respective designs. For notational convenience we take* $\sigma^2=1.00$ *in much that follows, reinstating* $\sigma^2\neq 1.00$ *as required for clarity.*

Remark 1 The model $\{\boldsymbol{Y}=\beta_0\boldsymbol{1}_n+\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{\epsilon}\}$, linear in $\boldsymbol{X}\in\mathbb{F}_{n\times k}$, has the model matrix $\mathbf{X}=[\mathbf{1}_n,\boldsymbol{X}]$ of order $(n\times(k+1))$. A full second-order model with $k=2$ regressors is

$$\{Y_i=\beta_0+x_{i1}\beta_1+x_{i2}\beta_2+x_{i1}^2\beta_{11}+x_{i2}^2\beta_{22}+x_{i1}x_{i2}\beta_{12}+\epsilon_i;1\le i\le n\}$$

so that rows of $\mathbf{X}$ $(n\times 6)$ are now $[1,x_1,x_2,x_1^2,x_2^2,x_1x_2]$.

2.3. *Ordered spaces*

Given a binary relation $\succeq_0$, a pair $(\mathcal{A},\succeq_0)$, is said to be *linearly ordered* if $\succeq_0$ is reflexive, transitive, antisymmetric and complete; *partially ordered* if reflexive, transitive and antisymmetric; and *preordered* if reflexive and transitive. A partially ordered set is a *lower semi-lattice* if for any two elements x,y in $\mathcal{A}$, there is a *glb* (greatest lower bound) $x\wedge y$ in $\mathcal{A}$; an *upper semi-lattice* if there is a *lub* (least upper bound) $x\vee y$ in $\mathcal{A}$; and a *lattice* if it is both a lower and upper semi-lattice. Spaces having lower and upper bounds are central to this study. Of note here are the following

2.3.1. *Bounds on* $(\mathbb{R}^k,\succeq_k)$

Clearly $(\mathbb{R}^k,\succeq_k)$ is a lattice with $\boldsymbol{a}\wedge\boldsymbol{b}=[a_1\wedge b_1,\ldots,a_k\wedge b_k]'$ and $\boldsymbol{a}\vee\boldsymbol{b}=[a_1\vee b_1,\ldots,a_k\vee b_k]'\}$, where $\{a_i\wedge b_i=\min(a_i,b_i)\}$ and $\{a_i\vee b_i=\max(a_i,b_i)\}$ for $\{1\le i\le k\}$. See,[3] for example.

2.3.2. *Bounds on* $(\mathbb{S}_k^+,\succeq_L)$

Ordered as in [1], $\boldsymbol{A}\succeq_L\boldsymbol{B}$ if and only if $\boldsymbol{A}-\boldsymbol{B}\in\mathbb{S}_k^0$, with $\boldsymbol{A}\succ_L\boldsymbol{B}$ whenever $\boldsymbol{A}-\boldsymbol{B}\in\mathbb{S}_k^+$. To motivate, let $\boldsymbol{\Sigma}$ and $\boldsymbol{\Omega}$ be dispersion matrices in $\mathbb{S}_k^+$ for Gauss–Markov estimators in two models, say A and B. The Rayleigh quotient here is

$$\gamma_k\le\frac{\boldsymbol{u}'\boldsymbol{\Sigma}\boldsymbol{u}}{\boldsymbol{u}'\boldsymbol{\Omega}\boldsymbol{u}}=\frac{\boldsymbol{w}'\boldsymbol{\Omega}^{-\frac{1}{2}}\boldsymbol{\Sigma}\,\boldsymbol{\Omega}^{-\frac{1}{2}}\boldsymbol{w}}{\boldsymbol{w}'\boldsymbol{w}}\le\gamma_1, \tag{1}$$

where $\{\boldsymbol{w}_i=\boldsymbol{\Omega}^{\frac{1}{2}}\boldsymbol{u}_i;1\le i\le k\}$, and $\{\gamma_1\ge\gamma_2\ge\cdots\ge\gamma_k>0\}$ are eigenvalues from the spectral decomposition $\boldsymbol{\Omega}^{-\frac{1}{2}}\boldsymbol{\Sigma}\,\boldsymbol{\Omega}^{-\frac{1}{2}}=\boldsymbol{Q}\,\mathrm{Diag}(\gamma_1,\ldots,\gamma_k)\,\boldsymbol{Q}'$ with $\boldsymbol{Q}\in\mathcal{O}(k)$. If neither $\boldsymbol{\Sigma}\succ\boldsymbol{\Omega}$ nor $\boldsymbol{\Omega}\succ\boldsymbol{\Sigma}$, then for some nonnegative (r,s) it would follow that

$$\{\gamma_1\ge\cdots\ge\gamma_r>1=\gamma_{r+1}=\cdots=\gamma_{r+s}>\gamma_{r+s+1}\ge\cdots\ge\gamma_k>0\}. \tag{2}$$

DEFINITION 2 *Regarding* $\boldsymbol{Q} = [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_k]$, *identify*

$$\begin{aligned} L_1 = S_p(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_r) &\iff \{\gamma_1, \ldots, \gamma_r\} \\ L_2 = S_p(\boldsymbol{q}_{r+1}, \ldots, \boldsymbol{q}_{r+s}) &\iff \{\gamma_{r+1}, \ldots, \gamma_{r+s}\} \\ L_3 = S_p(\boldsymbol{q}_{r+s+1}, \ldots, \boldsymbol{q}_k) &\iff \{\gamma_{r+s+1}, \ldots, \gamma_k\} \end{aligned}$$

as subspaces of $\mathbb{R}^k$ *corresponding to the indicated eigenvalues.*

Clearly at most two of $\{L_1, L_2, L_3\}$ may be empty, and $\boldsymbol{\Sigma} \succeq_L \boldsymbol{\Omega}$ asserts that L_3 is empty. For nonempty subspaces, the model A is uniformly more efficient than B for linear parametric functions in L_1; the models are equi-efficient for functions in L_2; and B is uniformly more efficient for functions in L_3. Here *efficiency* refers both to Fisher's variance comparisons in estimation (see [4]), and to Pitman's power comparisons in tests for linear hypotheses (see [5]).

That $(\mathbb{S}_k, \succeq_L)$ is not a lattice is noted in [6, p.142]. Nonetheless, $(\mathbb{S}_k^+, \succeq_L)$ admits lower and upper *spectral* bounds that are tight. This follows on imbedding $\boldsymbol{D}_\gamma = \text{Diag}(\gamma_1, \ldots, \gamma_k)$ into $(\mathbb{R}^k, \geq_k)$, itself a lattice as noted by Vulikh.[3] These spectral bounds emerge as follows; details are given in [2].

DEFINITION 3 *Given* $(\boldsymbol{\Sigma}, \boldsymbol{\Omega})$ *in* $(\mathbb{S}_k^+, \succeq_L)$ *such that neither* $\boldsymbol{\Sigma} \succ \boldsymbol{\Omega}$ *nor* $\boldsymbol{\Omega} \succ \boldsymbol{\Sigma}$. *Then the matrices given by*

$$\boldsymbol{\Sigma} \wedge \boldsymbol{\Omega} = \boldsymbol{\Omega}^{\frac{1}{2}} \boldsymbol{Q}(\boldsymbol{D}_\gamma \wedge \boldsymbol{I}_k)\boldsymbol{Q}' \boldsymbol{\Omega}^{\frac{1}{2}} \quad \text{and} \quad \boldsymbol{\Sigma} \vee \boldsymbol{\Omega} = \boldsymbol{\Omega}^{\frac{1}{2}} \boldsymbol{Q}(\boldsymbol{D}_\gamma \vee \boldsymbol{I}_k)\boldsymbol{Q}' \boldsymbol{\Omega}^{\frac{1}{2}}$$

are called the spectral glb and the spectral lub, respectively, for $(\boldsymbol{\Sigma}, \boldsymbol{\Omega})$.

Properties of these are essential to this study: principally, whether the operation $\boldsymbol{\Sigma} \wedge \boldsymbol{\Omega} = \boldsymbol{\Omega} \wedge \boldsymbol{\Sigma}$ commutes, and whether the spectral bounds are tight. Both are answered affirmatively, as summarized without proof in Theorem A.1 of the Appendix.

2.3.3. *Bounds on* $(\mathbb{F}_{n\times k}, \succeq)$

Ordering of $(\mathbb{F}_{n\times k}, \succeq)$ is induced from that of $(\mathbb{S}_k^+, \succeq_L)$ such that $\boldsymbol{X} \succeq \boldsymbol{Z}$ if and only if $\boldsymbol{X}'\boldsymbol{X} \succeq_L \boldsymbol{Z}'\boldsymbol{Z}$, requiring as before that $(\boldsymbol{X}, \boldsymbol{Z}) \in \mathbb{F}_{n\times k}$ should be of rank $k \leq n$. For further details see.[7] This ordering is invariant in that $\boldsymbol{X} \succeq \boldsymbol{Z}$ if and only if $\boldsymbol{PXB} \succeq \boldsymbol{QZB}$ for any $(\boldsymbol{P}, \boldsymbol{Q}) \in \mathcal{O}(n)$ and $\boldsymbol{B} \in \mathcal{G}l(k)$, and the antisymmetry axiom holds up to equivalence under $\mathcal{O}(n)$ acting from the left.

To proceed, begin with $(\boldsymbol{X}, \boldsymbol{Z}) \in \mathbb{F}_{n\times k}$; take

$$(\boldsymbol{X}, \boldsymbol{Z}) \to (\boldsymbol{X}(\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}, \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}) \to (\boldsymbol{P}\boldsymbol{D}_\lambda \boldsymbol{Q}', \boldsymbol{U}\boldsymbol{I}_k \boldsymbol{Q}'), \tag{3}$$

the latter as their singular decompositions, where $(\boldsymbol{P}, \boldsymbol{U}) \in \mathbb{H}_{n\times k}$ are the *left singular vectors,* $\boldsymbol{Q} \in \mathcal{O}(k)$ the *right singular vectors,* and $\boldsymbol{D}_\lambda = \text{Diag}(\lambda_1, \ldots, \lambda_k)$, the *singular values* of $\boldsymbol{X}(\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}$ ordered as $\{\lambda_1 \geq \cdots \geq \lambda_k > 0\}$. Again $(\mathbb{F}_{n\times k}, \succeq)$ is not a lattice, nor can it inherit lattice properties through its induced ordering. Nonetheless, $(\mathbb{F}_{n\times k}, \succeq)$ admits lower and upper *singular* bounds that themselves are tight. Details are supplied in [2], culminating in the following.

DEFINITION 4 *For* $(\boldsymbol{X}, \boldsymbol{Z}) \in (\mathbb{F}_{n\times k}, \succeq)$ *such that neither* $\boldsymbol{X} \succ \boldsymbol{Z}$ *nor* $\boldsymbol{Z} \succ \boldsymbol{X}$, *the matrices given by*

$$\boldsymbol{X} \wedge \boldsymbol{Z} = \boldsymbol{P}(\boldsymbol{D}_\lambda \wedge \boldsymbol{I}_k)\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}} \quad \text{and} \quad \boldsymbol{X} \vee \boldsymbol{Z} = \boldsymbol{P}(\boldsymbol{D}_\lambda \vee \boldsymbol{I}_k)\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}} \tag{4}$$

are called the singular glb and the singular lub, respectively, for $(\boldsymbol{X}, \boldsymbol{Z})$.

Properties of these are summarized without proof in Theorem A.2 of the Appendix. In addition, the following facts are used subsequently.

LEMMA 1 *Given* $(\boldsymbol{P}\boldsymbol{D}_\lambda\boldsymbol{Q}', \boldsymbol{U}\boldsymbol{I}_k\boldsymbol{Q}')$ *as the singular decompositions of Equation* (3) *with* $(\boldsymbol{P}, \boldsymbol{U}) \in \mathbb{H}_{n\times k}$, $\boldsymbol{Q} \in \mathcal{O}(k)$, *and* $\boldsymbol{D}_\lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_k)$; *the following hold.*

(i) $\boldsymbol{X}$ *may be recovered as* $\boldsymbol{X} = \boldsymbol{P}\boldsymbol{D}_\lambda\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}}$;
(ii) $\boldsymbol{Z}$ *may be recovered as* $\boldsymbol{Z} = \boldsymbol{P}\boldsymbol{I}_k\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}}$.

Proof Details are given in [2]; but it is crucial to note that the singular values of $\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}$ are repeated values of unity, so that its left and right singular vectors may be taken to be $(\boldsymbol{P}, \boldsymbol{Q})$, as for $\boldsymbol{X}(\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}$. ■

Remark 2 The principal driver, namely $\boldsymbol{X} \vee \boldsymbol{Z} = \boldsymbol{P}(\boldsymbol{D}_\lambda \vee \boldsymbol{I}_k)\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}}$, depends on $\boldsymbol{Z}$ only through $\boldsymbol{Z}'\boldsymbol{Z}$. Accordingly, $\boldsymbol{X}$ and $\boldsymbol{Z}$ need not have the same number of rows. For if $\boldsymbol{Z} \in \mathbb{F}_{s\times k}$, then the asserted properties continue to hold with $\boldsymbol{X} \vee \boldsymbol{Z} \in \mathbb{F}_{n\times k}$, while $\boldsymbol{Z} \vee \boldsymbol{X} \in \mathbb{F}_{s\times k}$ and $\boldsymbol{U} \in \mathbb{H}_{s\times k}$. For this reason the commutative property of Theorem A.1(iii) fails for the entity $\boldsymbol{X} \vee \boldsymbol{Z}$.

2.4. *Efficiency indices*

A real-valued function $\phi(\cdot)$ on $(\mathbb{S}_k^+, \succeq_L)$ is said to be *order-preserving* if $\boldsymbol{A} \succeq_L \boldsymbol{B}$ implies $\phi(\boldsymbol{A}) \geq \phi(\boldsymbol{B})$ on $\mathbb{R}^1$. Denote these by $\Phi(\mathbb{S}_k^+, \succeq_L)$, and by $\Phi_0(\mathbb{S}_k, \succeq_L)$ the subclass invariant under orthogonal congruence and thus depending only on the eigenvalues of its matrix argument. Functions in $\Phi(\mathbb{S}_k^+, \succeq_L)$ are characterized in [8].

Central to $\{\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$ are the Design $\boldsymbol{X}$, the Fisher Information matrix $\sigma^2\mathcal{I}(\hat{\boldsymbol{\beta}}) = \boldsymbol{X}'\boldsymbol{X}$, its inverse $\mathrm{V}(\hat{\boldsymbol{\beta}}) = \sigma^2\boldsymbol{\Sigma} = \sigma^2[\boldsymbol{\Sigma}_{ij}]$ with eigenvalues $\{Ch_i(\boldsymbol{\Sigma}) = \gamma_i; 1 \leq i \leq k\}$, and correlation matrix $\boldsymbol{R}$. Designs often are evaluated on the variances of estimation and prediction, to include the trace, determinant, and largest eigenvalue of $\boldsymbol{\Sigma}$ as the A, D and E indices in $\Phi_0(\mathbb{S}_k^+, \succeq_L)$ as in [9]. Earlier uses of determinants trace to $|\boldsymbol{R}|$ as the *scatter coefficient*, and to $|\boldsymbol{\Sigma}|$ as the *generalized variance,* as in Frisch [10] and Wilks.[11] Wald [12] initially proposed E-efficiencies, but opted instead for determinants, in his study of local power in the analysis of variance. Standard references include Federov,[13] Silvey,[14] Pukelsheim,[15] for example. Various concepts of *universal optimality* are reviewed and reassessed in [16] together with further references.

Nonetheless, the concept of *vector efficiency* is genuinely multidimensional, beyond capture by any scalar, as seen through axiomatics in [4]. Even Kiefer [17] advocated that competing designs be screened through multiple criteria. We survey a collection of scalar criteria (C_r) in current vogue, as summarized in Table 1, to include gages of efficiency for subsets of parameters. In regard to D-efficiency, recall that $\{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \leq \sigma^2 c_\alpha\}$ is the confidence ellipsoid for $\boldsymbol{\beta}$ with coefficient $1 - \alpha$ whenever c_α is the upper quantile of $\chi^2(\nu, 0)$ for suitable choice of ν. Excluding MV, Table 1 criteria all belong to Φ_0.

Notions of C_r*-optimality* are pervasive. If *minimal* is deemed *optimal,* then $\boldsymbol{X}_0$ is said to be C_r*-optimal* in a class $\{\boldsymbol{X}_i; i \in \mathscr{I}\}$, if and only if $\{C_r(\boldsymbol{X}_0) \leq C_r(\boldsymbol{X}_i); i \in \mathscr{I}\}$. For example, the design $\boldsymbol{X}_0$ with dispersion $\boldsymbol{\Sigma}_0$ is D-optimal if and only if $|\boldsymbol{\Sigma}_0| \leq |\boldsymbol{\Sigma}_i|$ for every $\{\boldsymbol{X}_i; i \in \mathscr{I}\}$.

Subset efficiencies are of continuing interest.[18–27] Similarly, if $\boldsymbol{\beta}' = [\boldsymbol{\beta}_1', \boldsymbol{\beta}_2']$; and if $\hat{\boldsymbol{\beta}}' = [\hat{\boldsymbol{\beta}}_1', \hat{\boldsymbol{\beta}}_2']$ is partitioned such that $\mathrm{V}(\hat{\boldsymbol{\beta}}_1) = \boldsymbol{\Sigma}_{11}$ and $\mathrm{V}(\hat{\boldsymbol{\beta}}_2) = \boldsymbol{\Sigma}_{22}$, then $\boldsymbol{X}_0$ is D_S-optimal for $\boldsymbol{\beta}_1$ if and only if $|\boldsymbol{\Sigma}_{011}| \leq |\boldsymbol{\Sigma}_{i11}|$ for every design in the class $\{\boldsymbol{X}_i; i \in \mathscr{I}\}$. Special properties of determinants support the following connection between D and D_S-efficiencies.

LEMMA 2 *Let* $\{\rho_1, \ldots, \rho_r\}$ *be Hotelling's* [28] *canonical correlations between vectors* $(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2)$ *of orders* $(r \times 1)$ *and* $(s \times 1)$ *with* $r \leq s$. *Then the D-efficiency of* $\hat{\boldsymbol{\beta}}$, *the* D_S^1*-efficiency of* $\hat{\boldsymbol{\beta}}_1$, *and*

Table 1. Efficiency criteria C_r for Design $\boldsymbol{X}$ having $\sigma^2\mathcal{I}(\hat{\boldsymbol{\beta}}) = \boldsymbol{X}'\boldsymbol{X}$ and $\mathrm{V}(\hat{\boldsymbol{\beta}}) = \boldsymbol{\Sigma} = [\boldsymbol{\Sigma}_{ij}]$, with eigenvalues $\{Ch_i(\boldsymbol{\Sigma}) = \gamma_i\}$ ordered as $\{\gamma_1 \geq \cdots \geq \gamma_k\}$, where C_{rS} refers to subset efficiencies, and with *minimization* as the operation yielding C_r-optimal designs.

Criteria	Description	Comments
A, A_S	$\mathrm{tr}(\boldsymbol{\Sigma})$, $\mathrm{tr}(\boldsymbol{\Sigma}_{11})$	A: Sum of $\mathrm{Var}(\hat{\beta}_i)$ of elements of $\hat{\boldsymbol{\beta}}$
D, D_S	$\lvert\boldsymbol{\Sigma}\rvert$, $\lvert\boldsymbol{\Sigma}_{11}\rvert$	$D^{\frac{1}{2}}$: $\propto \mathrm{Vol}\{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \leq \sigma^2 c_\alpha\}$
E, E_S	$Ch_1(\boldsymbol{\Sigma})$, $Ch_1(\boldsymbol{\Sigma}_{11})$	E: Maximal variance of $\boldsymbol{c}'\hat{\boldsymbol{\beta}}$, $\lVert\boldsymbol{c}\rVert = 1$
MV	$\max\{\mathrm{Var}(\hat{\beta}_i); 1 \leq i \leq k\}$	MV: Maximal variance of elements of $\hat{\boldsymbol{\beta}}$
T^{-1}	$1/\mathrm{tr}(\boldsymbol{X}'\boldsymbol{X})$	T: $\mathrm{tr}(\boldsymbol{X}'\boldsymbol{X}) = \sigma^2\,\mathrm{tr}(\mathcal{I}(\hat{\boldsymbol{\beta}}))$
$\boldsymbol{c}$	$\mathrm{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}})$	$\boldsymbol{c}'\boldsymbol{\beta}$: A distinguished linear function
I	$\sigma^2 \int_R \boldsymbol{u}'\boldsymbol{\Sigma}\boldsymbol{u}\,\mathrm{d}\boldsymbol{u}$	I: Integrated $\mathrm{Var}(\boldsymbol{u}'\hat{\boldsymbol{\beta}})$ over $R \subset \mathbb{R}^k$
V	$\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}$	V: Prediction variance at fixed $\boldsymbol{t} \in \mathbb{R}^k$

the D_S^2-efficiency of $\hat{\boldsymbol{\beta}}_2$, are related as $D = D_1 D_2 \times \prod_{i=1}^r (1 - \rho_i^2)$, where $\prod_{i=1}^r (1 - \rho_i^2)$ is the Alienation Coefficient of Hotelling.[28]

The following generalizes the notion of relative Fisher efficiency as the ratio of variances.

DEFINITION 5 *Identify $E_{\mathrm{ff}}{}^{C_r}(\boldsymbol{X} : \boldsymbol{Z}) = C_r(\boldsymbol{Z})/C_r(\boldsymbol{X})$ in $\mathbb{R}^1_+$ as the comparative efficiency of design $\boldsymbol{X}$ relative to $\boldsymbol{Z}$ with regard to the criterion C_r. Specifically, with dispersion matrices $\boldsymbol{\Sigma}$ from $\boldsymbol{X}$ and $\boldsymbol{\Omega}$ from $\boldsymbol{Z}$, then $E_{\mathrm{ff}}{}^{D}(\boldsymbol{X} : \boldsymbol{Z}) = \lvert\boldsymbol{\Omega}\rvert/\lvert\boldsymbol{\Sigma}\rvert$ is the ratio of their generalized variances.*

3. The principal findings

Given designs $\boldsymbol{X}$ and $\boldsymbol{Z}$, we examine essentials of $\boldsymbol{X}_M = \boldsymbol{X} \vee \boldsymbol{Z}$ as a design matrix itself, in comparison with $\boldsymbol{X}$ and $\boldsymbol{Z}$. Throughout we assume that neither $\boldsymbol{X} \succ \boldsymbol{Z}$ nor $\boldsymbol{Z} \succ \boldsymbol{X}$. For if $\boldsymbol{X} \succ \boldsymbol{Z}$ dominates strictly, then no extractable elements of $\boldsymbol{Z}$ are available in order to augment $\boldsymbol{X}$ to advantage, and similarly for $\boldsymbol{Z} \succ \boldsymbol{X}$. In what follows we often take $\boldsymbol{X}$ as reference, or otherwise $\boldsymbol{Z}$ as reference on reversing their roles. This section deals principally with $\boldsymbol{X} \in \mathbb{F}_{n\times k}$ and typically $\boldsymbol{Z} \in \mathbb{F}_{n\times k}$, or alternatively as $\boldsymbol{Z} \in \mathbb{F}_{s\times k}$. These on occasion may be required to be centered such that $\mathbf{1}_n'\boldsymbol{X} = \boldsymbol{0}$ and $\mathbf{1}_n'\boldsymbol{Z} = \boldsymbol{0}$.

3.1. *Matrix extremes*

Further properties of the spectral lower and upper bounds of Section 2.3.2, and of the singular lower and upper bounds of Section 2.3.3, are required. These are developed here in preparation for applications to experimental designs and linear inference. In regard to models $\{\boldsymbol{Y} = \beta_0\mathbf{1}_n + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$ and $\{\boldsymbol{Y} = \beta_0\mathbf{1}_n + \boldsymbol{Z}\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$, we suppose from the outset that these are *coherent* in the following sense.

DEFINITION 6 *Designs $\boldsymbol{X}$ and $\boldsymbol{Z}$ are said to be coherent when successive columns of $\boldsymbol{X}$ and $\boldsymbol{Z}$ are in the same units.*

First essentials for the maximal design $\boldsymbol{X}_M$, as induced through $(\boldsymbol{X}, \boldsymbol{Z}) \to \boldsymbol{X}_M = \boldsymbol{X} \vee \boldsymbol{Z}$, are listed in the following.

THEOREM 1 *Given $\boldsymbol{X}$ and $\boldsymbol{Z}$ as coherent designs; construct $\boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}$, $\boldsymbol{B} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}$ and $\boldsymbol{X}_M = \boldsymbol{X} \vee \boldsymbol{Z}$.*

(i) *The design $\boldsymbol{X}_M$ is coherent with $\boldsymbol{X}$ and $\boldsymbol{Z}$;*

(ii) *If the columns of* $\boldsymbol{X}$ *are centered such that* $\mathbf{1}_n'\boldsymbol{X} = \mathbf{0}$, *then the columns of* $\boldsymbol{X}_M$ *are centered and* $\mathbf{1}_n'\boldsymbol{X}_M = \mathbf{0}$;

(iii) *The matrix* $\boldsymbol{X}_M'\boldsymbol{X}_M$ *has the representation*

$$\boldsymbol{X}_M'\boldsymbol{X}_M = \boldsymbol{X}'\boldsymbol{P}\,\mathrm{Diag}(\boldsymbol{I}_r, \boldsymbol{I}_s, {\gamma_{r+s+1}}^{-1}, \ldots, {\gamma_k}^{-1})\boldsymbol{P}'\boldsymbol{X} \tag{5}$$

with (r, s, k) *and* $\boldsymbol{P}$ *as in Equation* (3);

(iv) *The matrices* $\boldsymbol{X}_M'\boldsymbol{X}_M$, $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are related as* $(\boldsymbol{X}_M'\boldsymbol{X}_M)^{-1} = \boldsymbol{A} \wedge \boldsymbol{B}$.

Proof (i) To identify units of columns of $\boldsymbol{X} \vee \boldsymbol{Z}$ in terms of $\boldsymbol{Z}$ and $\boldsymbol{X}$, infer from Lemma 1(ii) that $\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}} = \boldsymbol{I}_k\boldsymbol{P}'\boldsymbol{Z}$ and from Definition 3 that

$$\boldsymbol{X} \vee \boldsymbol{Z} = \boldsymbol{P}(\boldsymbol{D}_\lambda \vee \boldsymbol{I}_k)\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}} = \boldsymbol{G}'\boldsymbol{Z} \tag{6}$$

with $\boldsymbol{G}' = \boldsymbol{P}(\boldsymbol{D}_\lambda \vee \boldsymbol{I}_k)\boldsymbol{P}'$ of order $(n \times n)$ and rank k. That $\boldsymbol{G}'$ is dimensionless, acting from the left on columns of $\boldsymbol{Z}$, implies that units of $\boldsymbol{Z}$ are preserved column-wise into $\boldsymbol{X} \vee \boldsymbol{Z}$, giving conclusion (i). For conclusion (ii), Lemma 1(i) gives $\boldsymbol{X} = \boldsymbol{P}\boldsymbol{D}_\lambda\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}}$, so that $\mathbf{1}_n'\boldsymbol{X} = \mathbf{0} \Longleftrightarrow \mathbf{1}_n'\boldsymbol{P} = \mathbf{0}$ carries over to include $\mathbf{1}_n'(\boldsymbol{X} \vee \boldsymbol{Z}) = \mathbf{0}$ from Definition 3. To continue, apply $\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}} = {\boldsymbol{D}_\lambda}^{-1}\boldsymbol{P}'\boldsymbol{X}$ from Lemma 1(i) in going from the first to second lines of

$$\begin{aligned}
\boldsymbol{X}_M'\boldsymbol{X}_M &= (\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}}\boldsymbol{Q}(\boldsymbol{D}_\gamma \vee \boldsymbol{I}_k)\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}} && (7)\\
&= \boldsymbol{X}'\boldsymbol{P}{\boldsymbol{D}_\lambda}^{-1}(\boldsymbol{D}_\gamma \vee \boldsymbol{I}_k){\boldsymbol{D}_\lambda}^{-1}\boldsymbol{P}'\boldsymbol{X} && (8)\\
&= \boldsymbol{X}'\boldsymbol{P}\,\mathrm{Diag}(\boldsymbol{I}_r, \boldsymbol{I}_s, {\gamma_{r+s+1}}^{-1}, \ldots, {\gamma_k}^{-1})\boldsymbol{P}'\boldsymbol{X} && (9)
\end{aligned}$$

to give conclusion (iii), where $\boldsymbol{D}_\lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_k)$ are singular values as square roots of solutions of $|\boldsymbol{X}'\boldsymbol{X} - \gamma\boldsymbol{Z}'\boldsymbol{Z}| = 0$. For conclusion (iv), invert (7) as (a) and invoke Definition 2 as (b), where

$$\begin{aligned}
&(a)\ (\boldsymbol{X}_M'\boldsymbol{X}_M)^{-1} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}\boldsymbol{Q}(\boldsymbol{D}_\gamma \vee \boldsymbol{I}_k)^{-1}\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}\\
&(b)\ \boldsymbol{A} \wedge \boldsymbol{B} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}\boldsymbol{Q}(\boldsymbol{D}_\theta \wedge \boldsymbol{I}_k)\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}
\end{aligned}$$

such that elements of $\boldsymbol{D}_\theta$ are eigenvalues of $(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}}(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}}$, reciprocal to those of $(\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}$, so that $\boldsymbol{D}_\theta = {\boldsymbol{D}_\gamma}^{-1}$. That (a) and (b) are equal follows on noting that $(\boldsymbol{D}_\gamma \vee \boldsymbol{I}_k)^{-1} = ({\boldsymbol{D}_\gamma}^{-1} \wedge \boldsymbol{I}_k)$, to give conclusion (iv). ■

3.2. *First-order models*

Consider $\{\boldsymbol{Y} = \beta_0\mathbf{1}_n + f(\boldsymbol{X})\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$ such that $f(\boldsymbol{X}) = \boldsymbol{X}$. Throughout we take design matrices to be centered. Accordingly, the model matrix is $\mathbf{X} = [\mathbf{1}_n, \boldsymbol{X}]$; the Fisher Information matrix is $\sigma^2\mathcal{I}(\boldsymbol{\beta}) = \mathrm{Diag}(n, \boldsymbol{X}'\boldsymbol{X})$; consequently it suffices to consider only the design $\boldsymbol{X}$ and parameters $\boldsymbol{\beta} \in \mathbb{R}^k$, and similarly for $\boldsymbol{Z}$ and $\boldsymbol{X}_M = \boldsymbol{X} \vee \boldsymbol{Z}$. Basic properties follow on partitioning $\boldsymbol{X}'\boldsymbol{P} = \boldsymbol{F} = [\boldsymbol{F}_1, \boldsymbol{F}_2, \boldsymbol{F}_3] \in \mathbb{F}_{k\times r} \otimes \mathbb{F}_{k\times s} \otimes \mathbb{F}_{k\times t}$, with $r + s + t = k$, to the following effect.

Definition 7 *Let* $\boldsymbol{\theta} = \boldsymbol{F}'\boldsymbol{\beta} = [\boldsymbol{\theta}_1', \boldsymbol{\theta}_2', \boldsymbol{\theta}_3']'$ *as the canonical parameters in* $\mathbb{R}^r \otimes \mathbb{R}^s \otimes \mathbb{R}^t$, *and identify* $L_1 = S_p(\boldsymbol{F}_1)$, $L_2 = S_p(\boldsymbol{F}_2)$, *and* $L_3 = S_p(\boldsymbol{F}_3)$ *as the corresponding canonical subspaces.*

Fundamentals, to include extremal properties of Fisher Information and dispersion matrices for $\boldsymbol{X}_M$, are listed in the following.

THEOREM 2 *Consider designs* $\boldsymbol{X}$, $\boldsymbol{Z}$, *and* $\boldsymbol{X}_M = \boldsymbol{X} \vee \boldsymbol{Z}$; *their* OLS *solutions* $\hat{\boldsymbol{\beta}}_X$, $\hat{\boldsymbol{\beta}}_Z$, *and* $\hat{\boldsymbol{\beta}}_M$; *and* $\boldsymbol{\Sigma}$, $\boldsymbol{\Omega}$, *and* $\boldsymbol{\Xi}$ *as inverses of their Fisher Information matrices* $\mathcal{I}(\hat{\boldsymbol{\beta}}_X)$, $\mathcal{I}(\hat{\boldsymbol{\beta}}_Z)$ *and* $\mathcal{I}(\hat{\boldsymbol{\beta}}_M)$. *Then*

(i) $\mathcal{I}(\hat{\boldsymbol{\beta}}_M)$ *is the lub of* $(\mathcal{I}(\boldsymbol{X}), \mathcal{I}(\boldsymbol{Z}))$ *in* $(\mathbb{S}_k^+, \succeq_L)$, *i.e.* $\mathcal{I}(\hat{\boldsymbol{\beta}}_X) \vee \mathcal{I}(\hat{\boldsymbol{\beta}}_Z) = \mathcal{I}(\hat{\boldsymbol{\beta}}_M)$;
(ii) $\boldsymbol{\Xi}$ *is the glb of* $(\boldsymbol{\Sigma}, \boldsymbol{\Omega})$ *in* $(\mathbb{S}_k^+, \succeq_L)$, *i.e.* $\boldsymbol{\Sigma} \wedge \boldsymbol{\Omega} = \boldsymbol{\Xi}$;
(iii) *Under Assumptions A,* $\mathcal{L}(\hat{\boldsymbol{\beta}}_M) = N_k(\boldsymbol{\beta}, \sigma^2 \boldsymbol{\Xi})$;
(iv) $\boldsymbol{\Xi} = (\boldsymbol{P}'\boldsymbol{X})^{-1}\text{Diag}(\boldsymbol{I}_r, \boldsymbol{I}_s, \gamma_{r+s+1}, \ldots, \gamma_k)(\boldsymbol{X}'\boldsymbol{P})^{-1}$ *with* $\{1 > \gamma_{r+s+1} \geq \cdots \geq \gamma_k > 0\}$;
(v) $\boldsymbol{\Xi} \preceq_L \boldsymbol{\Sigma}$ *and* $\boldsymbol{\Xi} \preceq_L \boldsymbol{\Omega}$;
(vi) *Linear parametric functions in* L_3 *are estimated more precisely under* $\boldsymbol{X}_M$ *than* $\boldsymbol{X}$, *specifically,* $\text{V}(\hat{\boldsymbol{\theta}}_3) = \sigma^2\text{Diag}(\gamma_{r+s+1}, \ldots, \gamma_k)$ *under* $\boldsymbol{X}_M$ *and is* $\sigma^2\boldsymbol{I}_t$ *under* $\boldsymbol{X}$;
(vii) *Linear parametric functions in* L_1 *and* L_2 *are estimated with equal precision under* $\boldsymbol{X}_M$ *and* $\boldsymbol{X}$.
(viii) *Each efficiency criterion of Table 1 is no greater under* $\boldsymbol{X}_M$ *than under* $\boldsymbol{X}$ *or* $\boldsymbol{Z}$.

Proof Definition 3 gives $\boldsymbol{X}'\boldsymbol{X} \vee \boldsymbol{Z}'\boldsymbol{Z} = (\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}}\boldsymbol{Q}(\boldsymbol{D}_\gamma \vee \boldsymbol{I}_k)\boldsymbol{Q}'(\boldsymbol{Z}'\boldsymbol{Z})^{\frac{1}{2}}$, which is precisely (7), to give conclusion (i). Conclusion (ii) is a restatement of Theorem 1(iv). Conclusion (iii) follows since $\{\boldsymbol{Y} = \beta_0\boldsymbol{1}_n + \boldsymbol{X}_M\boldsymbol{\beta} + \boldsymbol{\epsilon}\}$ is a Gauss–Markov model. Conclusion (iv) follows on inverting (9), where $\boldsymbol{P}'\boldsymbol{X}$ is invertible by Lemma 1(i). Conclusion (v) follows from (ii) and Appendix Theorem A.1(i). Given $\boldsymbol{\theta} = \boldsymbol{F}'\boldsymbol{\beta}$ and $\boldsymbol{\theta}_3 = \boldsymbol{F}_3'\boldsymbol{\beta}$ from Definition 7, it follows that $\text{V}(\hat{\boldsymbol{\theta}}) = \sigma^2\text{Diag}(\boldsymbol{I}_r, \boldsymbol{I}_s, \gamma_{r+s+1}, \ldots, \gamma_k)$ under $\boldsymbol{X}_M$, whereas $\text{V}(\hat{\boldsymbol{\theta}}) = \sigma^2\boldsymbol{I}_k$ under $\boldsymbol{X}$. Linear coefficients in $L_3 = S_p(\boldsymbol{F}_3)$ are of type $\{\boldsymbol{F}_3\boldsymbol{c};\ \boldsymbol{c} \in \mathbb{R}^t\}$ so that $\boldsymbol{c}'\boldsymbol{\theta}_3 = \boldsymbol{c}'\boldsymbol{F}_3'\boldsymbol{\beta}$ is estimated with variance $\text{V}(\boldsymbol{c}'\boldsymbol{F}_3'\hat{\boldsymbol{\beta}}_M) = \sigma^2\boldsymbol{c}'\text{Diag}(\gamma_{r+s+1}, \ldots, \gamma_k)\boldsymbol{c}$ under $\boldsymbol{X}_M$, and with variance $\sigma^2\boldsymbol{c}'\boldsymbol{I}_t\boldsymbol{c}$ under $\boldsymbol{X}$, for every $\{\boldsymbol{c} \in S_p(\boldsymbol{F}_3)\}$, to give conclusion (vi). Conclusion (vii) follows similarly, Conclusion (viii) holds since $\boldsymbol{X}_M'\boldsymbol{X}_M \succeq_L \boldsymbol{X}'\boldsymbol{X}$ and $\boldsymbol{\Xi} \preceq_L \boldsymbol{\Sigma}$, which implies $\{Ch_i(\boldsymbol{\Xi}) \leq Ch_i(\boldsymbol{\Sigma}); 1 \leq i \leq k\}$ using Theorem 3 of Bellman,[29, p. 115] in regard to functions in $\Phi_0(\mathbb{S}_k^+, \succeq_L)$ monotone increasing in the ordered eigenvalues. ■

Remark 3 Through $\boldsymbol{X}_M = \boldsymbol{X} \vee \boldsymbol{Z}$, the orderings of Definitions 3 and 4 come together. Conclusion (i) asserts that $\mathcal{I}(\hat{\boldsymbol{\beta}}_M) = \max(\mathcal{I}(\hat{\boldsymbol{\beta}}_X), \mathcal{I}(\hat{\boldsymbol{\beta}}_Z))$, and conclusion (ii) that $\boldsymbol{\Xi} = \min(\boldsymbol{\Sigma}, \boldsymbol{\Omega})$.

4. Case studies

To fix ideas, we turn to examples where $\boldsymbol{X}$ and $\boldsymbol{Z}$ need not have the same number of rows, as noted in Remark 1.

4.1. *Single regressor* $k = 1$

Consider a straight-line model $\{Y_i = \beta_0 + x_i\beta_1 + \epsilon_i;\ 1 \leq i \leq n\}$ together with a prospective design having $n = 8$ centered regressors as in the first row of Table 2. Limited resources may dictate using only half the runs, considered provisionally as $\boldsymbol{z}'$ in the second row. It remains to see whether evidence in $\boldsymbol{x}$ might combine with that of $\boldsymbol{z}$ in order to complement the latter. The revised design vector $\boldsymbol{z}_M = \boldsymbol{z} \vee \boldsymbol{x}$ of Definition 4, also centered and of order $n = 4$, is listed in the third row of Table 2. Straight-line regression analyses follow routinely on noting that the first-order Fisher Information matrices at $\sigma^2 = 1.00$ are $\text{Diag}(8, 8.00)$, $\text{Diag}(4, 6.00)$, and $\text{Diag}(4, 8.00)$, respectively. Our tools effect improvements to first-order models as noted. For effects of higher order, consider in addition $\{Y_i = \beta_0 + x_i\beta_1 + x_i^2\beta_2 + \epsilon_i\}$, and similarly for $\boldsymbol{z}$ and $\boldsymbol{z}_M$, on appending squared elements to the rows of Table 2. Accordingly, as in Definition 1, take

Table 2. Regressor vectors $\boldsymbol{x}'(1\times 8)$, $z'(1\times 4)$, and $(z\vee \boldsymbol{x})' = z_M'(1\times 4)$ with $\alpha=\sqrt{2}$.

$\boldsymbol{x}'$	−1.0000	1.0000	$-\alpha$	α	−1.00	1.00	0.00	0.00
z'	−1.0000	1.0000	$-\alpha$	α				
z_M'	−1.1547	1.1547	−1.6330	1.6330				

Table 3. Summary from dispersion matrices $\boldsymbol{\Sigma}$, $\boldsymbol{\Omega}$ and $\boldsymbol{\Xi}$ corresponding to second-order models $\mathbf{X}$, $\mathbf{Z}$, and $\mathbf{Z}_M$, to include variances, eigenvalues $[\gamma_1\geq\gamma_2\geq\gamma_3]$, and efficiency diagnostics corresponding to A, D, E efficiency indices.

	Variances			Eigenvalues		
Item	$\boldsymbol{X}$	$\boldsymbol{Z}$	$\boldsymbol{Z}_M$	$\boldsymbol{X}$	$\boldsymbol{Z}$	$\boldsymbol{Z}_M$
$\widehat{\beta}_0$	0.3750	2.5000	2.5000	0.5702	3.4271	3.0159
$\widehat{\beta}_1$	0.1250	0.1667	0.1250	0.1250	0.1667	0.1250
$\widehat{\beta}_2$	0.2500	1.0000	0.5625	0.0548	0.0729	0.0466
	tr($\boldsymbol{\Gamma}$)			$10^2\lvert\boldsymbol{\Gamma}\rvert$		
	0.7500	3.6667	3.1875	0.3906	4.1648	1.7568
	Efficiency ratios					
	$E_{\rm ff}{}^A(\boldsymbol{Z}_M:\boldsymbol{Z})=1.1503$		$E_{\rm ff}{}^D(\boldsymbol{Z}_M:\boldsymbol{Z})=2.3707$		$E_{\rm ff}{}^E(\boldsymbol{Z}_M:\boldsymbol{Z})=1.1350$	

the model matrix $\mathbf{X}(8\times 3)$ to have have the typical rows $\{[1, x_i, x_i^2]; 1\leq i\leq 8\}$, and similarly for z and z_M, giving model matrices $\mathbf{Z}$ and $\mathbf{Z}_M$ of orders (4×3). Then $\mathbf{X}'\mathbf{X}$, $\mathbf{Z}'\mathbf{Z}$, and $\mathbf{Z}_M'\mathbf{Z}_M$ are their Fisher Information matrices, and $\boldsymbol{\Sigma}$, $\boldsymbol{\Omega}$, and $\boldsymbol{\Xi}$ their respective inverses. Diagonal elements of these are listed as variances in Table 3; also listed are their vectors of Eigenvalues; Traces as A-efficiencies; Determinants $10^2|\boldsymbol{\Sigma}| = \text{Det} = 0.3906$, $10^2|\boldsymbol{\Omega}| = 4.1648$, and $10^2|\boldsymbol{\Xi}| = 1.7568$. as D-efficiencies; together with efficiency ratios as in Definition 5.

Not listed are eigenvectors corresponding to largest eigenvalues. Their elements, $[0.7882, 0.0000, -0.6154]$ for $\boldsymbol{\Sigma}$, determine coefficients of normalized linear functions having the maximal variance $\gamma_1\sigma^2$ as in the following

$$\begin{aligned}
&\text{Var}(0.7882\widehat{\beta}_0 - 0.6154\widehat{\beta}_2) = 0.5702\sigma^2 \text{ for the model } \mathbf{X}\\
&\text{Var}(0.8506\widehat{\beta}_0 - 0.5257\widehat{\beta}_2) = 3.4271\sigma^2 \text{ for the model } \mathbf{Z}, \quad \text{and}\\
&\text{Var}(0.9090\widehat{\beta}_0 - 0.4168\widehat{\beta}_2) = 3.0159\sigma^2 \text{ for the model } \mathbf{Z}_M
\end{aligned} \tag{10}$$

and similarly for $\boldsymbol{\Omega}$ and $\boldsymbol{\Xi}$

Together with variances of $(\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2)$ apart from σ^2, these collectively display effects of combining evidence contained in $\boldsymbol{X}$ and $\boldsymbol{Z}$ into $\boldsymbol{Z}_M$, despite that $\boldsymbol{Z}$ itself is relatively uninformative owing in part to its small sample size. Moreover, improvement in first-order models is evident in Table 3, where $\text{Var}(\widehat{\beta}_1)$ ranges from $0.1250\,\sigma^2$ for model $\mathbf{X}$ with $n=8$, to $0.1667\,\sigma^2$ for $\mathbf{Z}$ and to $0.1250\,\sigma^2$ for the amended model $\mathbf{Z}_M$, both with $n=4$. Even for the parsimonious models with $n=4$, the variance $\text{Var}(\widehat{\beta}_2)$ has been reduced from $1.00\,\sigma^2$ under the half model $\mathbf{Z}$, to $0.5625\,\sigma^2$ under the modified half model $\mathbf{Z}_M$. These comparisons illustrate how evidence in $\boldsymbol{X}(8\times 3)$ combines to enhanced effect into $\boldsymbol{Z}_M(4\times 3)$. Diverse values for $\text{Var}(\widehat{\beta}_0)$ are attributed in part to different sample sizes. The trade-off here is more precise estimation of β_1 and β_2 with half the amended observations using $\boldsymbol{Z}_M$ instead of $\boldsymbol{Z}$, yet retaining the same $\text{Var}(\widehat{\beta}_0) = 2.50\,\sigma^2$ under $\boldsymbol{Z}_M$ and $\boldsymbol{Z}$. This comparison of $\boldsymbol{Z}$ with $\boldsymbol{Z}_M$ when $n=4$ is compelling when $\boldsymbol{X}$ with $n=8$ is out of reach owing to expense or to severely limited experimental material.

Table 4. Regressor vectors for designs $\boldsymbol{X}'(2 \times 9)$(CCD); $\boldsymbol{Z}'(2 \times 7)$(SCD); together with $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$ of order (2×7), where $\boldsymbol{\alpha} = \sqrt{2}$, $\boldsymbol{\gamma} = 1.70710$ and $\boldsymbol{\delta} = 0.29289$.

CCD	−1.00	−1.00	$-\boldsymbol{\alpha}$	$\boldsymbol{\alpha}$	0.00	0.00	1.00	1.00	0.00
$\boldsymbol{X}'$	−1.00	1.00	0.00	0.00	$-\boldsymbol{\alpha}$	$\boldsymbol{\alpha}$	− 1.00	1.00	0.00
SCD	−1.00	$-\boldsymbol{\alpha}$	$\boldsymbol{\alpha}$	0.00	0.00	1.00	0.00		
$\boldsymbol{Z}'$	−1.00	0.00	0.00	$-\boldsymbol{\alpha}$	$\boldsymbol{\alpha}$	1.00	0.00		
$(\boldsymbol{Z} \vee \boldsymbol{X})'$	−1.00	$-\boldsymbol{\gamma}$	$\boldsymbol{\gamma}$	$\boldsymbol{\delta}$	$-\boldsymbol{\delta}$	1.00	0.00		
$\boldsymbol{Z}_M'$	−1.00	$\boldsymbol{\delta}$	$-\boldsymbol{\delta}$	$-\boldsymbol{\gamma}$	$\boldsymbol{\gamma}$	1.00	0.00		

Table 5. Variances of OLS solutions at $\sigma^2 = 1.00$ under designs CCD($\boldsymbol{X}$), the SCD ($\boldsymbol{Z}$), together with $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$.

Design	β_0	β_1	β_2	β_{11}	β_{22}	β_{12}	Trace
$\boldsymbol{X}$	1.0000	0.1250	0.1250	0.3438	0.3438	0.2500	2.1876
$\boldsymbol{Z}$	1.0000	0.1875	0.1875	0.3750	0.3750	0.7500	2.8750
$\boldsymbol{Z}_M$	1.0000	0.1250	0.1250	0.1953	0.1953	0.4063	2.0469

Table 6. Eigenvalues $\{\gamma_1, \ldots, \gamma_6\}$ and determinants $(\text{Det} \times 10^{-4})$ of dispersion matrices of OLS solutions under designs CCD($\boldsymbol{X}$), SCD($\boldsymbol{Z}$), and $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$, together with the relative efficiencies $E_{\text{ff}}^{C_r}(\boldsymbol{Z}_M : \boldsymbol{Z}) = C_r(\boldsymbol{Z})/C_r(\boldsymbol{Z}_M)$ of Definition 5.

	Eigenvalues						
Design	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	Det
$\boldsymbol{X}$	1.5214	0.2500	0.1250	0.1250	0.1250	0.0411	0.3052
$\boldsymbol{Z}$	1.5587	0.7639	0.2500	0.1250	0.1250	0.0525	2.4416
$\boldsymbol{Z}_M$	1.3804	0.3185	0.1250	0.1250	0.0625	0.0355	0.1526
	Efficiency ratios						
	$E_{\text{ff}}^{A}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 1.4046$		$E_{\text{ff}}^{D}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 16.0000$			$E_{\text{ff}}^{E}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 1.1292$	

4.2. *CCD and SCD designs, $k = 2$*

The full second-order model for $k = 2$ regressors is

$$\{Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{11} x_{i1}^2 + \beta_{22} x_{i2}^2 + \beta_{12} x_{i1} x_{i2} + \epsilon_i\}. \tag{11}$$

as in Remark 1. The Central Composite Design (CCD) of Box and Wilson,[30] with axial $\alpha = \sqrt{2}$, has regressors as in rows 1 and 2 of Table 4. The Small Composite Design (SCD) of Hartley [31] with $\alpha = \sqrt{2}$, is in rows 3 and 4 of Table 4. The CCD is considered to be a 'gold standard', while the SCD is hailed as among the smallest of second-order designs, having one degree-of-freedom for 'error'.

Owing to established needs for minimal designs, we seek to augment the SCD using our tools, drawing from the CCD to possible advantage. Accordingly, $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$ is constructed as given in rows 5 and 6 of Table 4.

To continue, denote by $\mathbf{X}$, $\mathbf{Z}$, and $\mathbf{Z}_M$ the model matrices for the respective designs $\boldsymbol{X}$, $\boldsymbol{Z}$, and $\boldsymbol{Z}_M$, where the rows of $\mathbf{X}$ correspond to $[1, x_{i1}, x_{i2}, x_{i1}^2, x_{i2}^2, x_{i1}x_{i2}]$ as in Remark 1, and similarly for $\mathbf{Z}$ and $\mathbf{Z}_M$. To these ends, variances for OLS solutions are reported in Table 5, together with their sums as Trace for the *A*-efficiency indices. In addition, eigenvalues for the respective dispersion matrices are given in Table 6, together with their scaled determinants for *D*-efficiency, and the ratios $E_{\text{ff}}^{C_r}(\boldsymbol{Z}_M : \boldsymbol{Z}) = C_r(\boldsymbol{Z})/C_r(\boldsymbol{Z}_M)$ of Definition 5.

Table 7. Eigenvectors $\boldsymbol{q}_i' = [q_{11}, q_{12}, q_{13}, q_{14}, q_{15}, q_{16}]$ corresponding to largest eigenvalues γ_1 of dispersion matrices for OLS solutions under designs $\boldsymbol{X}$, $\boldsymbol{Z}$, and $\boldsymbol{Z}_M$.

Design	$[q_{11}$	q_{12}	q_{13}	q_{14}	q_{15}	$q_{16}]$	γ_1
$\boldsymbol{X}$	[0.8048	0.0000	0.0000	−0.4197	−0.4197	0.0000]	1.5214
$\boldsymbol{Z}$	[0.7776	0.0000	0.0000	−0.4344	−0.4344	0.1343]	1.5587
$\boldsymbol{Z}_M$	[0.8373	0.0000	0.0000	−0.3284	−0.3284	−0.2886]	1.3804

Of primal concern here is whether to replace the time-honored SCD with $\boldsymbol{Z}_M$. By most conventional bases for comparison, we see that $\boldsymbol{Z}_M$ dominates $\boldsymbol{Z}$ in having uniformly smaller variances than $\boldsymbol{Z}$, as well as in its uniformly smaller (A, D, E) efficiency indices.

In addition, from Table 6 the eigenvalues are smaller for $\boldsymbol{Z}_M$ than $\boldsymbol{Z}$ except γ_4, where they are equal. In consequence, for all efficiency indices in Φ_0 based on the eigenvalues of dispersion matrices where smaller is better, the SCD, as modified here into $\mathcal{Z}_M$, is at least as 'good' as Hartley's SCD.

To illustrate subset diagnostics, let $\boldsymbol{\theta}' = [\beta_0, \boldsymbol{\beta}']$, to be partitioned as $[\boldsymbol{\theta}_1', \boldsymbol{\theta}_2']'$ with $\boldsymbol{\theta}_1' = [\beta_0, \beta_1, \beta_2]$ and $\boldsymbol{\theta}_2' = [\beta_{11}, \beta_{22}, \beta_{12}]$, the respective first and second-order coefficients. As in Table 1 let A_1 and A_2 be the A_S criteria for $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, and similarly D_1 and D_2 for their D_S criteria. Then $A_1 = 1.2500$ and $A_2 = 0.7969$ from Table 5, with $A_1 + A_2 = 2.0469$. On partitioning $\boldsymbol{\Xi}$ to have principal blocks $\boldsymbol{\Xi}_{11}$ and $\boldsymbol{\Xi}_{22}$ of orders (3×3), we find $D_1 = |\boldsymbol{\Xi}_{11}| = 1.5625 \times 10^{-2}$ and $D_2 = |\boldsymbol{\Xi}_{22}| = 0.6836 \times 10^{-2}$. Since $D = 0.1526 \times 10^{-4}$ from Table 6, we have the decomposition $D = D_1 D_2 \times 0.1429$ as in Lemma 2, where $0.1429 = \prod_{i=1}^{3}(1 - \rho_i^2)$ is the *Coefficient of Alienation* of Hotelling [28] between $[\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2]$ and $[\widehat{\beta_{11}}, \widehat{\beta_{22}}, \widehat{\beta_{12}}]$.

Eigenvectors corresponding to largest eigenvalues of the dispersion matrices are listed in Table 7. These support a complete assessment in regard to E-optimality. As in Equation (10) from Table 3, we have for the CCD($\boldsymbol{X}$) design from Table 7 that

$$\mathrm{Var}(0.8048\widehat{\beta}_0 - 0.4197(\widehat{\beta_{11}} + \widehat{\beta_{22}})) = 1.5214\sigma^2,$$

and similarly for other designs as listed.

To proceed, we examine further consequences of the augmentation $\boldsymbol{Z} \to \boldsymbol{Z}_M$. Specifically, with $\mathrm{V}(\hat{\boldsymbol{\beta}}(\boldsymbol{Z})) = \boldsymbol{\Omega}$ and $\mathrm{V}(\hat{\boldsymbol{\beta}}(\boldsymbol{Z}_M)) = \boldsymbol{\Xi}$, we return to Definition 2 and undertake the spectral decomposition of $\boldsymbol{\Xi}^{-1/2}\boldsymbol{\Omega}\boldsymbol{\Xi}^{-1/2}$, independently of σ^2, recovering eigenvalues and eigenvectors as listed in Table 8. Accordingly, the subspaces are $L_1 = S_p(\boldsymbol{q}_i, \boldsymbol{q}_2, \boldsymbol{q}_3)$, $L_2 = S_p(\boldsymbol{q}_4, \boldsymbol{q}_5)$, and $L_3 = S_p(\boldsymbol{q}_6)$ from Definition 2. Here $\boldsymbol{Z}_M$ is more efficient than $\boldsymbol{Z}$ for linear parametric functions in L_1; equal efficiency applies for functions in L_2, to include $\beta_1 + \beta_2$; whereas design $\boldsymbol{Z}$ is superior to $\boldsymbol{Z}_M$ for $11.99\beta_0 + \beta_{11} + \beta_{22} + 3.24\beta_{12}$ on scaling $\boldsymbol{q}_6$. From columns 2 and 3, the relative efficiency $E_{\mathrm{ff}}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 2.00$ for inferences regarding $\beta_1 - \beta_2$ and $\beta_{11} - \beta_{22}$.

4.3. SCD *and Hexagon designs,* $k = 2$

This is a continuation of Section 4.2, but using a design smaller than the CCD for augmenting the SCD. The alternative is the Hexagon design with one center run as in [32, p.429], correcting a sign in the first row. Both designs are centered. The second-order model (11) continues to apply. Computations are carried out in parallel to those of Section 4.2, and are combined into Table 9, where efficiency ratios are as in Definition 5. The SCD design $\boldsymbol{Z}$ and its values are those reported in Section 4.2 without the need for further repetition.

Our principal focus here surrounds the viability of replacing the time-honored SCD with $\boldsymbol{Z}_M$. By most conventional bases for comparison, we see that $\boldsymbol{Z}_M$ dominates $\boldsymbol{Z}$ in having uniformly smaller variances than $\boldsymbol{Z}$, as well as in its uniformly smaller (A, D, E) efficiency indices.

Table 8. Eigenvalues and eigenvectors of $\boldsymbol{\Xi}^{-1/2}\boldsymbol{\Omega}\boldsymbol{\Xi}^{-1/2}$.

Eigenvalues					
γ_1	γ_2	γ_3	γ_4	γ_5	γ_6
5.8117	2.0000	2.0000	1.0000	1.0000	0.6883
Eigenvectors					
$\boldsymbol{q}_i$	$\boldsymbol{q}_2$	$\boldsymbol{q}_3$	$\boldsymbol{q}_4$	$\boldsymbol{q}_5$	$\boldsymbol{q}_6$
−0.0525	0.0000	0.0000	−0.2779	0.0000	−0.9592
0.0000	0.7071	0.0000	0.0000	−0.7071	0.0000
0.0000	−0.7071	0.0000	0.0000	−0.7071	0.0000
−0.5865	0.0000	0.7071	0.3868	0.0000	−0.0800
−0.5865	0.0000	−0.7071	0.3868	0.0000	−0.0800
0.5562	0.0000	0.0000	0.7896	0.0000	−0.2592

Table 9. Regressor vectors for designs $\boldsymbol{H}$ (Hexagon) and $\boldsymbol{Z}_M^H = \boldsymbol{Z} \vee \boldsymbol{H}$; variances of OLS solutions at $\sigma^2 = 1.00$; eigenvalues for dispersion matrices; traces and determinants (Det $\times 10^{-4}$) of dispersion matrices of OLS solutions; and relative efficiency ratios $E_{\text{ff}}^{C_r}(\boldsymbol{Z}_M^H : \boldsymbol{Z}) = C_r(\boldsymbol{Z})/C_r(\boldsymbol{Z}_M^H)$.

Desn	Regressors						
$\boldsymbol{H}'$	− 0.7071	− 1.4142	− 0.7071	0.7071	1.4142	0.7071	0.0000
	− 1.2247	0.0000	1.2247	1.2247	0.0000	− 1.2247	0.0000
$\boldsymbol{Z}_M^{H'}$	− 1.0000	− 1.5313	− 0.1589	1.0000	1.5713	0.1589	0.0000
	− 1.0000	0.1589	1.5731	1.0000	− 0.1589	− 1.5731	0.0000
	Variances						
Desn	β_0	β_1	β_2	β_{11}	β_{22}	β_{12}	Trace
$\boldsymbol{H}$	1.0000	0.1667	0.1667	0.3750	0.3750	0.3333	2.4167
$\boldsymbol{Z}_M^H$	1.0000	0.1458	0.1458	0.2465	0.2465	0.4861	2.2707
	Eigenvalues						
Desn	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	Det
$\boldsymbol{H}$	1.5288	0.3333	0.1667	0.1667	0.1667	0.0545	1.2864
$\boldsymbol{Z}_M^H$	1.3986	0.4534	0.1667	0.1250	0.0833	0.0438	0.4821
	Efficiency ratios						
	$E_{\text{ff}}^A(\boldsymbol{Z}_M^H : \boldsymbol{Z}) = 1.0643$		$E_{\text{ff}}^D(\boldsymbol{Z}_M^H : \boldsymbol{Z}) = 2.6683$		$E_{\text{ff}}^E(\boldsymbol{Z}_M^H : \boldsymbol{Z}) = 1.0931$		

In addition, from Table 9 the eigenvalues are smaller for $\boldsymbol{Z}_M$ than $\boldsymbol{Z}$ except γ_3, where they are equal. In consequence, for all efficiency indices based on the eigenvalues of dispersion matrices where smaller is better, the SCD, as modified here into $\mathcal{Z}_M$, is at least as 'good' as Hartley's SCD.

4.4. *Improving optimal designs, $k = 2$*

Apart from *universal* optimality, conventional designs declared to be *optimal* are configured on a single criterion for optimality. Nonetheless, views among users often differ markedly as to criteria appropriate even to a particular experiment. The tools presented here enable us to merge two designs, separately optimal under one of two criteria, into a single

Table 10. Variances and traces for OLS solutions at $\sigma^2 = 1.00$, in the D-optimal design $\boldsymbol{X}$, the I-optimal design $\boldsymbol{Z}$, and $\boldsymbol{X}_M = \boldsymbol{X} \vee \boldsymbol{Z}$, together with eigenvalues, determinants (Det $\times 10^{-5}$), and efficiency ratios for each design.

Variances							
Design	β_0	β_1	β_2	β_{11}	β_{22}	β_{12}	Trace
$\boldsymbol{X}$	0.2993	0.0693	0.0693	0.2820	0.2820	0.0833	1.0852
$\boldsymbol{Z}$	0.1786	0.0833	0.0833	0.2143	0.2143	0.1250	0.8988
$\boldsymbol{Z}_M$	0.1786	0.0669	0.0669	0.1381	0.1381	0.0805	0.6691
Eigenvalues							
Design	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	Det
$\boldsymbol{X}$	0.5043	0.3414	0.0853	0.0666	0.0654	0.0222	0.1418
$\boldsymbol{Z}$	0.3301	0.2500	0.1250	0.0833	0.0833	0.0270	0.1933
$\boldsymbol{Z}_M$	0.2725	0.1611	0.0805	0.0671	0.0667	0.0211	0.0334
Efficiency ratios							
	$E^A_{\text{ff}}(\boldsymbol{Z}_M : \boldsymbol{X}) = 1.6219$		$E^D_{\text{ff}}(\boldsymbol{Z}_M : \boldsymbol{X}) = 4.2455$		$E^E_{\text{ff}}(\boldsymbol{Z}_M : \boldsymbol{X}) = 1.8506$		
	$E^A_{\text{ff}}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 1.3433$		$E^D_{\text{ff}}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 5.7874$		$E^E_{\text{ff}}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 1.2114$		

Table 11. Eigenvectors $\boldsymbol{q}_i{}' = [q_{11}, q_{12}, q_{13}, q_{14}, q_{15}, q_{16}]$ corresponding to largest eigenvalues γ_1 of dispersion matrices of OLS solutions under designs $\boldsymbol{X}$, $\boldsymbol{Z}$, and $\boldsymbol{Z}_M$.

Design	$[q_{11}$	q_{12}	q_{13}	q_{14}	q_{15}	$q_{16}]$	γ_1
$\boldsymbol{X}$	[0.7573	−0.0704	0.0704	−0.4564	−0.4564	0.0010]	0.5043
$\boldsymbol{Z}$	[−0.7071	0.0000	0.0000	0.5000	0.5000	0.0000]	0.3301
$\boldsymbol{Z}_M$	[0.7914	0.0000	0.0000	−0.4323	−0.4323	0.0047]	0.2725

nearby design dominating both designs on these and other criteria. This represents a mathematical first step towards constructing designs under *Compound Criteria* as set forth in [33, p.870 ff]. See also [26] and [27].

Accordingly consider designs with $k = 2$ regressors optimal for second-order response models. As given in [34], these are listed for completeness in Table 12, where Design $\boldsymbol{X}$, now centered, is D-optimal; Design $\boldsymbol{Z}$ is I-optimal; and $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$ is constructed as in Definition 4. Jones and Goos [34] focus on I-optimality as critical in dose-response methodology, where prediction often dominates the need for estimation and hypothesis testing, both relevant to D-optimality. To illustrate the enhanced capacity of $\boldsymbol{Z}_M$ over $\boldsymbol{X}$ and $\boldsymbol{Z}$, we undertake the following analyses, taking model matrices to have $n = 20$ rows as $[1, x_1, x_2, x_1^2, x_2^2, x_1x_2]$ from $\boldsymbol{X}$, and similarly for $\boldsymbol{Z}$ and $\boldsymbol{Z}_M$, as in Remark 1, We examine prospects for replacing the separately optimal $\boldsymbol{X}$ and $\boldsymbol{Z}$ with $\boldsymbol{Z}_M$. Specifically, summary computations are reported in Table 10, to include variances, eigenvalues, traces and determinants of the (6×6) dispersion matrices $(\boldsymbol{\Sigma}, \boldsymbol{\Omega}, \boldsymbol{\Xi})$ from $(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Z}_M)$. By conventional bases for comparison, $\boldsymbol{Z}_M$ dominates both $\boldsymbol{X}$ and $\boldsymbol{Z}$ in having uniformly smaller variances excluding $\text{Var}(\widehat{\beta_0})$, as well as uniformly smaller (A, D, E) efficiency indices. In addition, the eigenvalues are uniformly smaller for $\boldsymbol{Z}_M$ than $\boldsymbol{Z}$, and smaller than $\boldsymbol{X}$ except γ_4 and γ_5.

Eigenvectors corresponding to the largest eigenvalue γ_1 are reported in Table 11, in order to characterize the workings of E-optimality. As before, these vectors identify those standardized linear functions having the maximal variance $\gamma_1\sigma^2$. For example, for $\boldsymbol{Z}_M$: $\text{Var}(0.7914\widehat{\beta_0} - 0.4323(\widehat{\beta_{11}} + \widehat{\beta_{22}}) + 0.0047\widehat{\beta_{12}}) = 0.2725\sigma^2$.

To evaluate the I-criterion for each design, Borkowski [35, p.75] has noted that, with symbolic software such as Maple, the I-criterion can be evaluated numerically. For example, with regard

Table 12. Design values for the D-optimal $\boldsymbol{X}(20 \times 2)$, and for the I-optimal $\boldsymbol{Z}(20 \times 2)$, together with $(\boldsymbol{Z} \vee \boldsymbol{X}) = \boldsymbol{Z}_M(20 \times 2)$.

$\boldsymbol{X}$		$\boldsymbol{Z}$		$\boldsymbol{Z}_M$	
− 1.05	1.05	− 1.00	1.00	− 1.11430	1.11430
− 1.05	1.05	− 1.00	1.00	− 1.11430	1.11430
− 1.05	1.05	0.00	1.00	0.00187	1.11617
− 0.05	1.05	0.00	1.00	0.00187	1.11617
0.95	1.05	1.00	1.00	1.11803	1.11803
0.95	1.05	1.00	1.00	1.11803	1.11803
0.95	1.05	− 1.00	0.00	− 1.11617	− 0.00187
− 1.05	0.05	− 1.00	0.00	− 1.11617	− 0.00187
− 0.05	0.05	0.00	0.00	0.00000	0.00000
− 0.05	0.05	0.00	0.00	0.00000	0.00000
0.95	0.05	0.00	0.00	0.00000	0.00000
0.95	0.05	0.00	0.00	0.00000	0.00000
− 1.05	− 0.95	1.00	0.00	1.11617	0.00187
− 1.05	− 0.95	1.00	0.00	1.11617	0.00187
− 1.05	− 0.95	− 1.00	− 1.00	− 1.11803	− 1.11803
− 0.05	− 0.95	− 1.00	− 1.00	− 1.11803	− 1.11803
− 0.05	− 0.95	0.00	− 1.00	− 0.00187	− 1.11617
0.95	− 0.95	0.00	− 1.00	− 0.00187	− 1.11617
0.95	− 0.95	1.00	− 1.00	1.11430	− 1.11430
0.95	− 0.95	1.00	− 1.00	1.11430	− 1.11430

Table 13. Values for $I(\boldsymbol{\Sigma})$, $I(\boldsymbol{\Omega})$, and $I(\boldsymbol{\Xi})$ for varying regions $[-b, b] \times [-b, b]$.

b	$I(\boldsymbol{\Sigma})$	$I(\boldsymbol{\Omega})$	$I(\boldsymbol{\Xi})$
0.6667	0.2425	0.1579	0.1591
1.0000	0.2321	0.1829	0.1676
1.1180	0.2493	0.2127	0.1833
1.5000	0.4546	0.4462	0.3200

to $\boldsymbol{u}' = [1, x_1, x_2, x_1^2, x_2^2, x_1x_2]$ and $\boldsymbol{\Sigma}$, predictive variance is the expected value for $\boldsymbol{u}'\boldsymbol{\Sigma}\boldsymbol{u}$ over the region of interest, namely $R = \{[x_1, x_2] \in [-b, b] \times [-b, b]\}$ such that the required iterated and standardized integral, from Table 1 at $\sigma^2 = 1.00$, takes the form

$$I(\boldsymbol{\Sigma}) = \frac{1}{(2b)^2} \int_{-b}^{b} \int_{-b}^{b} (\boldsymbol{u}'\boldsymbol{\Sigma}\boldsymbol{u})\, \mathrm{d}u_1\, \mathrm{d}u_2,$$

and similarly for $I(\boldsymbol{\Omega})$ and $I(\boldsymbol{\Xi})$. Table 13 reports these integral values for varying b. In particular, the I-optimal design takes the value $I(\boldsymbol{\Omega}) = 0.1829$ on $[-1, 1] \times [-1, 1]$; corresponding values are $I(\boldsymbol{\Sigma}) = 0.2321$ and $I(\Xi) = 0.1676$.

To continue, we return to Definition 2 and undertake the spectral decomposition of $\boldsymbol{\Xi}^{-1/2}\boldsymbol{\Omega}\boldsymbol{\Xi}^{-1/2}$, again independently of σ^2, recovering eigenvalues and eigenvectors as listed in Table 14. Accordingly, the subspaces of Definition 2 are $L_1 = S_p(\boldsymbol{q}_i, \boldsymbol{q}_2, \boldsymbol{q}_3, \boldsymbol{q}_4, \boldsymbol{q}_5)$. L_2 is empty, and $L_3 = S_p(\boldsymbol{q}_6)$. Here $\boldsymbol{Z}_M$ is more efficient than $\boldsymbol{Z}$ for linear functions in L_1, where $\gamma_i > 1$, whereas design $\boldsymbol{Z}$ is superior to $\boldsymbol{Z}_M$ for $(-9.997\,\beta_0 + 0.185\,(\beta_{11} + \beta_{22}) - 0.014\beta_{12})$ with relative efficiency $1/0.8421 = 1.2134$, on scaling $\boldsymbol{q}_6$. Moreover, from column 2 the relative efficiency is $E_{\mathrm{ff}}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 1.5521$ for inferences regarding $\beta_{11} - \beta_{22}$, and from column 4 $E_{\mathrm{ff}}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 1.250$ for $\beta_1 + \beta_2$.

In summary, the fact of merging the D-optimal $\boldsymbol{X}$ and the I-optimal $\boldsymbol{Z}$ into $\boldsymbol{Z}_M$ improves not only their D and I efficiencies, but in addition $E_{\mathrm{ff}}{}^{C_r}(\boldsymbol{Z}_M, \boldsymbol{Z}) > 1.00$ for every efficiency criterion of Table 1, owing to the strictly ordered eigenvalues for $\boldsymbol{\Omega}$ and $\boldsymbol{\Xi}$. As noted, the foregoing developments represent a first step towards constructing *Compound Criteria* as set forth in

Table 14. Eigenvalues and eigenvectors of $\boldsymbol{\Xi}^{-1/2}\boldsymbol{\Omega}\boldsymbol{\Xi}^{-1/2}$.

Eigenvalues					
γ_1	γ_2	γ_3	γ_4	γ_5	γ_6
1.8840	1.5521	1.5515	1.2500	1.2417	0.8241
Eigenvectors					
$\boldsymbol{q}_i$	$\boldsymbol{q}_2$	$\boldsymbol{q}_3$	$\boldsymbol{q}_4$	$\boldsymbol{q}_5$	$\boldsymbol{q}_6$
0.0261	0.0000	−0.0025	0.0000	0.0000	−0.9997
0.0000	0.0000	0.0000	0.7071	0.7071	0.0000
0.0000	0.0000	0.0000	0.7071	−0.7071	0.0000
0.7063	0.7071	−0.0294	0.0000	0.0000	0.0185
0.7063	−0.7071	−0.0294	0.0000	0.0000	0.0185
0.0417	0.0000	0.9991	0.0000	0.0000	−0.0014

[33, p.870 ff] and, from another perspective, in Dette [26] and Dette and Franke.[27] The present approach could incorporate further optimality criteria iteratively on choosing yet another design optimal under a third criterion, then merging it in turn with $\boldsymbol{Z}_M$. Unfortunately, the foregoing analyses demand that we register the following.

Disclaimer 1 Properties of $\boldsymbol{Z}_M$ appear to contradict the claim of Jones and Goos [34] that $\boldsymbol{X}$ is D-optimal and $\boldsymbol{Z}$ is I-optimal. For the D-criterion is smaller for $\boldsymbol{Z}_M$ than $\boldsymbol{X}$ in excess by the factor 0.2355, and the I-criterion smaller for $\boldsymbol{Z}_M$ than $\boldsymbol{Z}$ by the factor 0.9163, so that both are dominated by $\boldsymbol{Z}_M$, seen in Table 12 to be nearby to $\boldsymbol{X}$ and $\boldsymbol{Z}$.

4.5. *H311B and CCD, $k = 3$*

The full second-order model in $k = 3$ regressors is

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_{11} x_{i1}^2 + \beta_{22} x_{i2}^2 + \beta_{33} x_{i3}^2 + \beta_{12} x_{i1} x_{i2} + \beta_{13} x_{i1} x_{i3} + \beta_{23} x_{i2} x_{i3} + \epsilon_i. \tag{12}$$

Owing to time constraints or limited experimental material, small second-order designs are critical and often at a premium. The CCD with one center run requires $n = 15$ runs per experiment. Among the smallest second-order designs for $k \geq 3$ are the hybrid designs of Roquemore.[36] Prominent among these is H311 B with center in $n = 11$ runs, 73.3% the size of the standard CCD in $n = 15$ runs. The design points for H311 B are listed in Table 15.

The CCD with axial point α is widely available, as in [32], for example, and not repeated here. Designating the CCD with $n = 15$ and $\alpha = 2.915476$ as design $\boldsymbol{X}$, We seek to recast H311 B as $\boldsymbol{Z}$ using the *CCD* as $\boldsymbol{X}$ through $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$ as reported in Table 15.

To examine properties of designs H311 B and $\boldsymbol{Z}_M$, we proceed as before, giving essential portions of the extensive output as reported in Table 16. Specifically, variances of OLS solutions, and eigenvalues of dispersion matrices for the CCD, H311 B, and $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$ designs; and traces and determinants of their dispersion matrices, are listed in Table 16. In particular, the critical comparisons are between $\boldsymbol{Z}$ and $\boldsymbol{Z}_M$, both of order (11×3). It is seen that variances are uniformly smaller for $\boldsymbol{Z}_M$ than $\boldsymbol{Z}$. as are their eigenvalues. In consequence, all efficiency indices of Table 1 are smaller for $\boldsymbol{Z}_M$ than $\boldsymbol{Z}$. Specifically, the (A, D, E) criteria for $\boldsymbol{Z}_M$ are $(1.2938, 0.1271, 1.0536)$, and $(1.4253, 3.6484, 1.0841)$ for $\boldsymbol{Z}$. Clearly $H311B$ may be supplanted by $\boldsymbol{Z}_M$ in $n = 11$ runs to enhanced effect of the latter.

Table 15. Design matrices for $\boldsymbol{Z}$ = H311 B and $\boldsymbol{Z}_M$ = H311 B $\vee$ CCD of order (11 × 3).

Design matrices					
$\boldsymbol{Z}$ = H311 B			$\boldsymbol{Z}_M$ = H311 B $\vee$ CCD		
0.0000	0.0000	2.4495	0.00000	− 0.00000	2.73701
0.0000	0.0000	− 2.4495	− 0.00000	0.00000	− 2.73701
− 0.7507	2.1063	1.0000	− 1.58118	1.93646	1.11902
2.1063	0.7507	1.0000	1.93646	1.58118	1.11902
0.7507	− 2.1063	1.0000	1.58118	− 1.93646	1.11902
− 2.1063	− 0.7507	1.0000	− 1.93646	− 1.58118	1.11902
0.7507	2.1063	− 1.0000	− 0.00000	2.50000	− 1.11902
2.1063	− 0.7507	− 1.0000	2.50000	0.00000	− 1.11902
− 0.7507	− 2.1063	− 1.0000	− 0.00000	− 2.50000	− 1.11902
− 2.1063	0.7507	− 1.0000	− 2.50000	0.00000	− 1.11902
0.0000	0.0000	0.0000	0.00000	0.00000	0.00000

Table 16. Variances of OLS solutions and eigenvalues of dispersion matrices for the CCD($\boldsymbol{X}$), H311 B($\boldsymbol{Z}$) and $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$ designs; and traces and determinants of these dispersion matrices of type $\boldsymbol{\Gamma}$.

Design characteristics						
Design	$\boldsymbol{X}$	$\boldsymbol{Z}$	$\boldsymbol{Z}_M$	$\boldsymbol{X}$	$\boldsymbol{Z}$	$\boldsymbol{Z}_M$
Estimates	Variances			Eigenvalues		
$\widehat{\beta_0}$	0.2582	1.0000	1.0000	0.2757	1.0841	1.0536
$\widehat{\beta_1}$	0.0400	0.0500	0.0400	0.1250	0.0501	0.0400
$\widehat{\beta_2}$	0.0400	0.0500	0.0400	0.1250	0.0500	0.0400
$\widehat{\beta_3}$	0.0400	0.0500	0.0400	0.1250	0.0500	0.0400
$\widehat{\beta_{11}}$	0.0123	0.0416	0.0277	0.0400	0.0500	0.0319
$\widehat{\beta_{22}}$	0.0123	0.0416	0.0277	0.0400	0.0500	0.0319
$\widehat{\beta_{33}}$	0.0123	0.0416	0.0267	0.0400	0.0500	0.0286
$\widehat{\beta_{12}}$	0.1250	0.0500	0.0277	0.0069	0.0168	0.0119
$\widehat{\beta_{13}}$	0.1250	0.0500	0.0319	0.0069	0.0167	0.0107
$\widehat{\beta_{23}}$	0.1250	0.0500	0.0319	0.0056	0.0077	0.0051
Item		tr($\boldsymbol{\Gamma}$)			$\lvert\boldsymbol{\Gamma}\rvert \times 10^{14}$	
	0.7901	1.4253	1.2938	0.9175	3.6484	0.1271

Designs $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{Z}_M$ are roughly comparable in variances apart from β_0 and $\{\beta_{ij};\, i \neq j\}$. But the relative efficiencies are $E_{\text{ff}}(\boldsymbol{Z} : \boldsymbol{X} | \widehat{\beta}_{ij}, i \neq j) = 2.50$. Moreover, $E_{\text{ff}}(\boldsymbol{Z}_M : \boldsymbol{X} | \widehat{\beta}_{ij}, i \neq j) \geq 3.9185$. These ratios are critical in dose-response experiments where $\{\beta_{ij} > 0\}$ designates *synergistic* factors (x_i, x_j), whereas $\{\beta_{ij} < 0\}$ characterizes *antagonistic* factors, sought on occasion to be estimated with greater precision. In this regard both $\boldsymbol{Z}$ and $\boldsymbol{Z}_M$ are superior to $\boldsymbol{X}$ despite its $n = 15$ runs, but with $\boldsymbol{Z}_M$ superior to $\boldsymbol{Z}$ by factors $E_{\text{ff}}(\boldsymbol{Z}_M : \boldsymbol{Z} | \widehat{\beta}_{ij}, i \neq j) \geq 1.5674$. In short, $\{\beta_{ij}; i \neq j\}$ are estimated in $\boldsymbol{Z}_M$ as if based on 64% of the observations in the H311B experiment, and 26% of the observations in the CCD.

4.6. *Improving optimal designs, k = 3*

The full second-order model in $k = 3$ regressors, given at the beginning of Section 4.5, is continued here, with parameters $[\beta_0, \boldsymbol{\beta}']$ and $\boldsymbol{\beta}$ as regression coefficients of order (9 × 1). Based on an experiment described in Exercise 11.6 of Box and Draper,[37] Gilmour and Trinca [38] give in their Table 2 two designs, namely Designs I and III with $k = 3$ and $n = 18$. Design I, to be labeled $\boldsymbol{X}$, is constructed to be D_S-optimal, and Design III, labeled $\boldsymbol{Z}$, to be A_S-optimal,

Table 17. Variances of OLS solutions and eigenvalues of dispersion matrices for the D_S-optimal design $\boldsymbol{X}$, the A_S-optimal $\boldsymbol{Z}$, and $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$; the traces and determinants of dispersion matrices $\boldsymbol{\Gamma}_{22}$ of order (9×9) excluding $\widehat{\beta}_0$; and relative efficiencies $E_{\text{ff}}{}^{C_r}(\boldsymbol{Z}_M : \boldsymbol{Z})$ and $E_{\text{ff}}{}^{C_r}(\boldsymbol{Z}_M : \boldsymbol{X})$ under D_S and A_S efficiency criteria.

	Design characteristics					
Design	$\boldsymbol{X}$	$\boldsymbol{Z}$	$\boldsymbol{Z}_M$	$\boldsymbol{X}$	$\boldsymbol{Z}$	$\boldsymbol{Z}_M$
$\hat{\boldsymbol{\beta}}$	Variances			Eigenvalues of $\boldsymbol{\Gamma}_{22}$		
$\widehat{\beta}_0$	0.4947	0.3318	0.3318			
$\widehat{\beta}_1$	0.0650	0.0686	0.0620	0.1928	0.1365	0.1182
$\widehat{\beta}_2$	0.0650	0.0686	0.0627	0.1430	0.1259	0.1046
$\widehat{\beta}_3$	0.0650	0.0686	0.0627	0.1044	0.1259	0.0984
$\widehat{\beta}_{11}$	0.0967	0.0817	0.0676	0.1044	0.0965	0.0868
$\widehat{\beta}_{22}$	0.0967	0.0817	0.0689	0.0644	0.0737	0.0655
$\widehat{\beta}_{33}$	0.0967	0.0817	0.0689	0.0644	0.0737	0.0622
$\widehat{\beta}_{12}$	0.1142	0.1156	0.0964	0.0570	0.0583	0.0571
$\widehat{\beta}_{13}$	0.1142	0.1156	0.0964	0.0488	0.0535	0.0451
$\widehat{\beta}_{23}$	0.1142	0.1156	0.0966	0.0488	0.0535	0.0444
Item		$\text{tr}(\boldsymbol{\Gamma}_{22})$			$\lvert\boldsymbol{\Gamma}_{22}\rvert \times 10^{11}$	
	0.8278	0.7975	0.6823	16.8804	18.8977	4.9170
	Efficiency ratios					
	$E_{\text{ff}}{}^{A_S}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 1.1688$ $E_{\text{ff}}{}^{A_S}(\boldsymbol{Z}_M : \boldsymbol{X}) = 1.2132$		$E_{\text{ff}}{}^{D_S}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 3.8433$ $E_{\text{ff}}{}^{D_S}(\boldsymbol{Z}_M : \boldsymbol{X}) = 3.4331$		$E_{\text{ff}}{}^{E_S}(\boldsymbol{Z}_M : \boldsymbol{Z}) = 1.1548$ $E_{\text{ff}}{}^{E_S}(\boldsymbol{Z}_M : \boldsymbol{X}) = 1.6311$	

where the subsets of parameters are elements of $\boldsymbol{\beta}$, thus excluding β_0. We seek an improved design by recasting $\boldsymbol{Z}$ into $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$ as in Definition 4, proceeding as in Section 4.2 for designs having $k = 2$. After centering, Table A1 gives the initial D_S-optimal design $\boldsymbol{X}$, and also the A_S-optimal design $\boldsymbol{Z}$, taken from [38]. This table is relegated to Appendix 2 for completeness. As constructed in Definition 4, $\boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$ is of order (18×3) and is listed in Table A1, also centered by Theorem 1(ii). To examine properties of designs $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{Z}_M$ we proceed as before, reporting essentials of the output in Table 17, to include variances of the *OLS* solutions. As both D_S and A_S refer to the subset $\boldsymbol{\beta}$ of $[\beta_0, \boldsymbol{\beta}']$, the relevant submatrix $\boldsymbol{\Gamma}_{22}(9 \times 9)$ of the typical dispersion matrix $\boldsymbol{\Gamma}(10 \times 10)$ is found on eliminating the first row and column of the latter. Accordingly, eigenvalues of $\boldsymbol{\Gamma}_{22}$ of order (9×1) are listed in Table 17; each A_S-criterion is the trace, and each D_S-criterion the determinant, of the corresponding submatrix $\boldsymbol{\Gamma}_{22}$.

It is seen that the D_S-criterion is smaller for design $\boldsymbol{X}$ than $\boldsymbol{Z}$, and the A_S-criterion smaller for $\boldsymbol{Z}$ than $\boldsymbol{X}$, as intended under their respective notions of optimality. In contrast, the $\boldsymbol{Z}_M$ design subsumes both, by a considerable margin in the case of D_S-efficiency. Moreover, variances are uniformly smaller for $\boldsymbol{Z}_M$ than either $\boldsymbol{X}$ or $\boldsymbol{Z}$, as are their E_S-criteria. In short, it is seen that the mapping $(\boldsymbol{X}, \boldsymbol{Z}) \to \boldsymbol{Z}_M = \boldsymbol{Z} \vee \boldsymbol{X}$ offers a substantial improvement in essentials over the D_S-optimal design $\boldsymbol{X}$ and the A_S-optimal $\boldsymbol{Z}$.

Disclaimer 2 As in Section 4.4, $\boldsymbol{Z}_M$ appears to refute the claim of Gilmour and Trinca [38] that $\boldsymbol{X}$ is D_S-optimal and $\boldsymbol{Z}$ is A_S-optimal. For the D_S-criterion is smaller for $\boldsymbol{Z}_M$ than $\boldsymbol{X}$ by the considerable factor 0.2913, and the A_S-criterion smaller for $\boldsymbol{Z}_M$ than $\boldsymbol{Z}$ by the factor 0.8555. Accordingly, $\boldsymbol{X}$ and $\boldsymbol{Z}$ are dominated in both criteria by $\boldsymbol{Z}_M$, seen in Table A1 to be nearby to $\boldsymbol{X}$ and $\boldsymbol{Z}$.

4.7. *A retrospective*

The analyses heretofore have emphasized the role of small designs, especially in industrial experiments as often advocated in the literature. On the other hand, Gilmour and Trinca [38] point

to needs for reliable estimates for σ^2, especially in regard to estimation and tests for hypotheses. On revisiting our several examples, we now offer further comments on this matter. It is pertinent that center runs in $\boldsymbol{Z}$ map into center runs in $\boldsymbol{Z}_M$ as seen in Tables 4, 9, 12 and 15. This is verified mathematically in developments surrounding Definition 4. Observe that the Fisher Information matrix and its inverse are invariant to the number of center runs, themselves supporting estimates for 'pure' error. Accordingly, two center runs may be added to the SCD of Table 4, now the same size as the CCD but having two degrees of freedom for 'pure' error. As the modified SCD, the comparative efficiencies of $\boldsymbol{Z}_M$ to the CCD remain as reported.

Similarly, center runs may be added to the SCD and Hexagon designs of Table 9 without altering essential comparisons given there. The modified I-optimal design $\boldsymbol{Z}_M$ of Table 12 not only dominates the D-optimal design $\boldsymbol{X}$ and the I-optimal $\boldsymbol{Z}$, but it has the advantage of four center runs. The H311 B design of Table 15 may be augmented to have four additional center runs, now the size of the CCD. Other comparisons of $\boldsymbol{Z}_M$ with the CCD remain intact as in Section 4.5, but now with four degrees of freedom for 'pure' error in design $\boldsymbol{Z}_M$ in contrast to the CCD.

5. Conclusions

Two design matrices $\mathbf{X}$ and $\mathbf{Z}$ of full rank are studied, to include their *singular least upper bound* $\boldsymbol{X}_M = \boldsymbol{X} \vee \boldsymbol{Z}$ as an enhanced design matrix. Both $\boldsymbol{X}$ and $\boldsymbol{Z}$ are centered, and Theorem 1(ii) assures that $\boldsymbol{X}_M = \boldsymbol{X} \vee \boldsymbol{Z}$ is also centered. The principal focus is that some singular values of $\boldsymbol{X}(\boldsymbol{Z}'\boldsymbol{Z})^{-1/2}$ are less than, and some greater than unity. Corresponding to these designs are the model matrices $\mathbf{X}$, $\mathbf{Z}$, and $\mathbf{X}_M$ together with $\boldsymbol{\Sigma}$, $\boldsymbol{\Omega}$ and $\boldsymbol{\Xi}$ as the dispersion matrices of the respective OLS estimators. Whereas $\boldsymbol{X} \vee \boldsymbol{Z}$ is maximal, the corresponding dispersion matrix is minimal, being the *spectral lower bound* $\boldsymbol{\Xi} = \boldsymbol{\Sigma} \wedge \boldsymbol{\Omega}$ as in Theorem 1(iv). The designs $\boldsymbol{X}$ and $\boldsymbol{Z}$ can differ in numbers of rows, allowing a smaller and parsimonious $\boldsymbol{Z}$ to be enhanced by a larger and more informative $\boldsymbol{X}$, so that $\boldsymbol{Z}_M$ is of the same order as $\boldsymbol{Z}$ and has improved efficiency criteria at least those of $\boldsymbol{Z}$. The success of our algorithm is that it formally identifies elements of $\boldsymbol{Z}$ complementary to those of $\boldsymbol{X}$, and combines these with $\boldsymbol{X}$ to form $\boldsymbol{X}_M$.

Case studies as reported treat second-order models in which the design matrix is augmented to include the constant and second-order terms. These demonstrate that A, D, E and other efficiencies are improved with $\boldsymbol{Z}_M$, often with a relatively small change from $\boldsymbol{Z}$ to $\boldsymbol{Z}_M$ as seen in Section 4.2. Moreover, subspaces in which linear parametric functions are estimated with greater, equal, or lesser precision between designs are identified in Definition 7.

It is shown in succession that the SCD, Hexagon, H311 B and other standard second-order designs may be improved through the operation $\boldsymbol{X} \vee \boldsymbol{Z}$. This in turn provides the methodology for users to explore yet other alternatives to a catalog of second-order and other experimental designs long available in the literature. In particular, our construction of $\boldsymbol{X} \vee \boldsymbol{Z}$ from designs optimal under separate criteria offers a first step towards invoking *Compound Criteria* in the spirit of Kiefer.[33]

We harbor grave concerns regarding claims in Sections 4.4 and 4.6 that pairs of designs are optimal; see Disclaimers listed there. Extenuating issues persist. In the Discussion following Gilmour and Trinca,[38] Bradley Jones and Dibyen Majumdar commented that 'The authors do not discuss the details of their algorithms except for saying that it is an exchange algorithm.' Yet Gilmour and Trinca [38] avow thereby to construct 'near optimum' designs. Jones and Goos [34] claim to have restricted their designs to the 9 points in the unit square. But this is refuted by points of $\boldsymbol{X}$ outside this grid as seen in Table 12. It remains unclear as to how the algorithms for finding D, D_S, A, A_S and I-optimal designs, as asserted, could have missed finding $\boldsymbol{X}_M$ or $\boldsymbol{Z}_M$ in each case, had these been allowed as candidates, and then to have delivered it as optimal instead.

Acknowledgements

Valuable comments followed the presentation of [2] to the *Joint AMS-IMS-SIAM Summer Research Conference on Stochastic Inequalities,* Seattle, Washington, July 1991. The late Professor Samuel Kotz remarked as most important the approximately three pages that served to set the stage for the present manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Loewner K . Über monotone Matrixfunktionen. Math Z. 1934;38:177–216.
[2] Jensen DR. Matrix extremes and related stochastic bounds. Stochastic inequalities. IMS Lecture Notes - Monograph Series, Vol. 22. 1993. p. 133–144.
[3] Vulikh BZ. Introduction to the theory of partially ordered spaces. Groningen: Wolters-Noordhoff; 1967.
[4] Jensen DR. Vector efficiency in multiparameter estimation. Linear Algebra Appl. 1991;151:143–155.
[5] Jensen DR, Ramirez DE. Efficiency comparisons in linear inference. J Stat Plann Infer. 1993;37:51–68.
[6] Halmos P. Finite-dimensional vector spaces. 2nd ed. New York, NY: Van Nostrand; 1958.
[7] Jensen DR. Invariant ordering and order preservation. In: Tong YL, editor. Inequalities in statistics and probability. IMS Lecture Notes/Monograph Series, Vol. 5. Hayward, CA; 1984.
[8] Marshall AW, Walkup DW, Wets RJB. Order preserving functions: applications to majorization and order statistics. Pacific J Math. 1967;23:569–584.
[9] Kiefer J. Optimum experimental designs. J Royal Statist Soc Ser B. 1959;21:272–304.
[10] Frisch R. Correlation and scatter in statistical variables. Nordic Statist J. 1929;1:36–102.
[11] Wilks SS. Certain generalizations in the analysis of variance. Biometrika. 1932;24:471–494.
[12] Wald A. On the efficient design of statistical investigations. Ann Math Statist. 1943;14:134–140.
[13] Fedorov VV. Theory of optimal experiments. Studden WJ, Klimko EM, translator. New York, NY: Academic Press; 1972.
[14] Silvey SD. Optimal design. London: Chapman and Hall; 1980.
[15] Pukelsheim F. Optimal design of experiments. New York, NY: Wiley; 1993.
[16] Druilhet P. Conditions for optimality in experimental designs. Linear Algebra Appl. 2004;388:147–157.
[17] Kiefer J. Optimum design: variation in structure and performance under change of criterion. Biometrika. 1975;62:277–288.
[18] Sibson R. D_A-optimality and duality. In: Gani J, Sarkadi K, Vincze I, editors. Progress in statistics, Proc. 9th European meeting of statisticians, Budapest, 1972, Vol. 2. Amsterdam: North-Holland; 1974.
[19] Studden WJ. D_s-optimal designs for polynomial regression using continued fractions. Ann Statist. 1980;8: 1132–1141.
[20] Huda S. On D_s-efficiency of some D-optimal designs. Calcutta Statist Assoc Bull. 1985;34:119–122.
[21] Huda S. On some D_s-optimal designs in spherical regions. Comm Statist Theory Methods. 1991;20:2965–2985.
[22] Huda S, Khan IH. On efficiency of some D-optimal designs in subset estimation. J Indian Soc Agricultural Statist. 1995;47:151–164.
[23] Lee CM-S. D-optimal designs for polynomial regression when lower degree parameters are more important. Utilitas Math. 1988;34:53–63.
[24] Lim YB, Studden WJ. Efficient D_s-optimal designs for multivariate polynomial regression on the q-cube. Ann Statist. 1988;16:1225–1240.
[25] Dette H. A generalization of D- and D_1-optimal designs in polynomial regression. Ann Statist. 1990;18:1784–1804.
[26] Dette H. On a mixture of the D- and D_1-optimality criterion in polynomial regression. J Statist Plan Inference. 1993;35:233–249.
[27] Dette H, Franke T. Robust designs for polynomial regression by maximizing a minimum of D- and D_1-efficiencies. Ann Statist. 2001;29:1024–1049.
[28] Hotelling H. Relations between two sets of variates. Biometrika. 1936;28:321–377.
[29] Bellman R. Introduction to matrix analysis. New York, NY: McGraw-Hill; 1960.
[30] Box GEP, Wilson KB. On the experimental attainment of optimum conditions. J Royal Statist Soc Ser B. 1951;13: 1–45.
[31] Hartley HO. Smallest composite design for quadratic response surfaces. Biometrics. 1959;15:611–624.
[32] Myers RH, Montgomery DC. Response surface methodology. 2nd ed. New York, NY: Wiley; 2002.
[33] Kiefer J. General equivalence theory for optimum designs (approximate theory). Ann Statist. 1974;2:849–879.
[34] Jones B, Goos P. I-optimal versus D-optimal split-plot response surface designs. J Qual Technol. 2012;44:85–101.
[35] Borkowski J. A comparison of a prediction variance criteria for response surface designs. J Qual Technol. 2003;35:70–77.
[36] Roquemore KG. Hybrid designs for quadratic response surfaces. Technometrics. 1976;18:419–423.
[37] Box GEP, Draper NR. Response surfaces, mixtures, and ridge analyses. 2nd ed. New York, NY: Wiley; 2007.
[38] Gilmour SG, Trinca LA. Optimum design of experiments for statistical inference. Appl Statist. 2012;61:345–401.

Appendix 1

The following theorems, stated without proof from [2], are reproduced here in that the original *Proceedings* source may not be widely available. Here $\Phi(\mathbb{S}_k^+, \succeq_L)$ is as in Section 2.3, and $\Phi(\mathbb{F}_{n\times k}, \succeq)$ the monotone functions induced as in [2].

THEOREM A.1 *Let* $\{\boldsymbol{A}\wedge\boldsymbol{B}, \boldsymbol{A}\vee\boldsymbol{B}\}$ *and* $\{\boldsymbol{B}\wedge\boldsymbol{A}, \boldsymbol{B}\vee\boldsymbol{A}\}$ *be spectral glb's and lub's as in Definition 1. Then for any* $(\boldsymbol{A},\boldsymbol{B})$ *in* $(\mathbb{S}_k^+, \succeq_L)$,

(i) $\boldsymbol{A}\wedge\boldsymbol{B} \preceq_L \{\boldsymbol{A},\boldsymbol{B}\} \preceq_L \boldsymbol{A}\vee\boldsymbol{B}$,
(ii) $\phi(\boldsymbol{A}\wedge\boldsymbol{B}) \le \{\phi(\boldsymbol{A}), \phi(\boldsymbol{B})\} \le \phi(\boldsymbol{A}\vee\boldsymbol{B})$ *for each* $\phi \in \Phi(\mathbb{S}_k^+, \succeq_L)$, *and*
(iii) $\boldsymbol{A}\wedge\boldsymbol{B} = \boldsymbol{B}\wedge\boldsymbol{A}$ *and* $\boldsymbol{A}\vee\boldsymbol{B} = \boldsymbol{B}\vee\boldsymbol{A}$.*Moreover, the bounds are tight in the sense that*
(iv) *if* $\{\boldsymbol{A},\boldsymbol{B}\} \preceq_L \boldsymbol{T}$ *and* $\boldsymbol{T} \preceq_L \boldsymbol{A}\vee\boldsymbol{B}$, *then* $\boldsymbol{T} = \boldsymbol{A}\vee\boldsymbol{B}$, *and*
(v) *if* $\{\boldsymbol{A},\boldsymbol{B}\}\succeq_L\boldsymbol{S}$ *and* $\boldsymbol{S}\succeq_L\boldsymbol{A}\wedge\boldsymbol{B}$, *then* $\boldsymbol{S} = \boldsymbol{A}\wedge\boldsymbol{B}$.

THEOREM A.2 *Given* $(\boldsymbol{X},\boldsymbol{Z}) \in (\mathbb{F}_{n\times k}, \succeq)$ *and* $\boldsymbol{X}\wedge\boldsymbol{Z}$ *and* $\boldsymbol{X}\vee\boldsymbol{Z}$ *as in Definition 2. Then*

(i) $\boldsymbol{X}\wedge\boldsymbol{Z} \preceq \{\boldsymbol{X},\boldsymbol{Z}\} \preceq \boldsymbol{X}\vee\boldsymbol{Z}$;
(ii) $\phi(\boldsymbol{X}\wedge\boldsymbol{Z}) \le \{\phi(\boldsymbol{X}), \phi(\boldsymbol{Z})\} \le \phi(\boldsymbol{X}\vee\boldsymbol{Z})$ *for each* $\phi \in \Phi(\mathbb{F}_{n\times k}, \succeq)$; *and*
(iii) $\boldsymbol{X}\wedge\boldsymbol{Z}$ *and* $\boldsymbol{X}\vee\boldsymbol{Z}$ *are determined up to equivalence under* $\mathcal{O}(n)$ *acting from the left. Moreover, the bounds are tight in the sense that*
(iv) *if* $\{\boldsymbol{X},\boldsymbol{Z}\} \preceq \boldsymbol{T}$ *and* $\boldsymbol{T} \preceq \boldsymbol{X}\vee\boldsymbol{Z}$, *then* $\boldsymbol{T}$ *is equivalent to* $\boldsymbol{X}\vee\boldsymbol{Z}$, *and*
(v) *if* $\{\boldsymbol{X},\boldsymbol{Z}\} \succeq \boldsymbol{S}$ *and* $\boldsymbol{S} \succeq \boldsymbol{X}\wedge\boldsymbol{Z}$, *then* $\boldsymbol{S}$ *is equivalent to* $\boldsymbol{X}\wedge\boldsymbol{Z}$.

Appendix 2

Table A1. Design matrix for the D_S-optimal design $\boldsymbol{X}$ of order (18×3), for the A_S-optimal $\boldsymbol{Z}$, and for $\boldsymbol{Z}_M = \boldsymbol{Z}\vee\boldsymbol{X}$ of orders (18×3).

Design matrices					
D_S-optimal $\boldsymbol{X}$			$\boldsymbol{X}$ Continued		
−1.0805	−0.9195	−0.9195	1.1442	−1.1442	0.0805
−1.0805	−0.9195	1.0805	−0.0805	0.0805	−1.6515
−1.0805	1.0805	−0.9195	−0.0805	0.0805	1.8126
−1.0805	1.0805	1.0805	−0.0805	−1.6515	0.0805
0.9195	−0.9195	1.0805	−0.0805	1.8126	0.0805
0.9195	1.0805	−0.9195	−1.8126	0.0805	0.0805
0.9195	1.0805	1.0805	1.6515	0.0805	0.0805
−0.0805	−1.1442	−1.1442	−0.0805	0.0805	0.0805
1.1442	0.0805	−1.1442	−0.0805	0.0805	0.0805
A_S-optimal $\boldsymbol{Z}$				$\mathbf{Z}_M = \mathbf{Z}\vee\mathbf{X}$	
−0.9445	−0.9444	−0.9444	−0.9354	−0.9500	−0.9500
−0.9445	−0.9444	−0.9444	−0.9354	−0.9500	−0.9500
−0.9445	−0.9444	1.0556	−0.9354	−0.9653	1.1391
−0.9445	1.0556	−0.9444	−0.9354	1.1391	−0.9653
−0.9445	1.0556	1.0556	−1.0554	1.1239	1.1238
1.0556	−0.9444	−0.9444	1.1654	−1.0121	−1.0121
1.0556	−0.9444	1.0556	1.1054	−1.0274	1.0771
1.0556	1.0556	−0.9444	1.1054	1.0771	−1.0274
1.0556	1.0556	1.0556	1.0454	1.0618	1.0618
−1.6765	0.0556	0.0556	−1.7643	0.1097	0.1097
1.7876	0.0556	0.0556	1.8743	0.0021	0.0021
0.0556	−1.6765	0.0556	0.1070	−1.7534	0.0691
0.0556	1.7876	0.0556	0.0031	1.8652	0.0427
0.0556	0.0556	−1.6765	0.1070	0.0691	−1.7534
0.0556	0.0556	1.7876	0.0031	0.0423	1.8652
0.0556	0.0556	0.0556	0.0550	0.0559	0.0559
0.0556	0.0556	0.0556	0.0550	0.0559	0.0559
0.0556	0.0556	0.0556	0.0550	0.0559	0.0559