A delimitation of the support of optimal designs for Kiefer's ϕ_p-class of criteria

L. Pronzato

Laboratoire I3S, CNRS/Université de Nice-Sophia Antipolis
Bât. Euclide, Les Algorithmes, 2000 route des Lucioles, BP 121
06903 Sophia Antipolis cedex, France
pronzato@i3s.unice.fr

March 21, 2013

Abstract

The paper extends the result of Harman and Pronzato [Stat. & Prob. Lett., 77:90–94, 2007], which corresponds to $p = 0$, to all strictly concave criteria in Kiefer's ϕ_p-class. Let ξ be any design on a compact set $\mathscr{X} \subset \mathbb{R}^m$ with a nonsingular information matrix $\mathbf{M}(\xi)$, and let δ be the maximum of the directional derivative $F_{\phi_p}(\xi, \mathbf{x})$ over all $\mathbf{x} \in \mathscr{X}$. We show that any support point $\mathbf{x}_*$ of a ϕ_p-optimal design satisfies the inequality $F_{\phi_p}(\xi, \mathbf{x}_*) \geq h_p[\mathbf{M}(\xi), \delta]$, where the bound $h_p[\mathbf{M}(\xi), \delta]$ is easily computed: it requires the determination of the unique root of a simple univariate equation (polynomial when p is integer) in a given interval. The construction can be used to accelerate algorithms for ϕ_p-optimal design and is illustrated on an example with A-optimal design.

keywords Approximate design; optimum design; support points; design algorithm
MSC 62K05; 90C46

1 Introduction and motivation

For $\mathscr{X}$ a compact subset of $\mathbb{R}^m$, denote by Ξ the set of design measures (i.e., probability measures) on $\mathscr{X}$ and by $\mathbf{M}(\xi)$ the information matrix

$$\mathbf{M}(\xi) = \int_{\mathscr{X}} \mathbf{x}\mathbf{x}^\top \, \xi(\mathrm{d}\mathbf{x}) .$$

We suppose that there exists a nonsingular design on $\mathscr{X}$ (i.e., there exists a $\xi \in \Xi$ such that $\mathbf{M}(\xi)$ is nonsingular) and we denote by Ξ^+ the set of such designs. We consider an optimal design problem on $\mathscr{X}$ defined by the maximization of a design criterion $\phi(\xi) = \Phi[\mathbf{M}(\xi)]$ with respect to $\xi \in \Xi$. One may refer to Pukelsheim (1993, Chap. 5) for a presentation of desirable properties that make a criterion $\Phi(\cdot)$ appropriate to measure the information provided by ξ. Here we shall focuss our attention on design criteria that correspond to the ϕ_p-class considered by Kiefer (1974). More precisely, we consider the positively homogeneous form of such criteria and, for any $\mathbf{M} \in \mathbb{M}$, the set of symmetric non-negative definite $m \times m$ matrices, we denote

$$\Phi_p^+(\mathbf{M}) = \left[\frac{1}{m}\,\mathrm{tr}(\mathbf{M}^{-p})\right]^{-1/p} , \tag{1}$$

with the continuous extension $\Phi_p^+(\mathbf{M}) = 0$ when $\mathbf{M}$ is singular and $p \geq 0$. A design measure ξ_p^* that maximizes $\phi_p(\xi) = \Phi_p^+[\mathbf{M}(\xi)]$ will be said ϕ_p-optimal. Note that when $p \neq 0$ the maximization of $\Phi_p^+(\mathbf{M})$ is equivalent to the minimization of $[\mathrm{tr}(\mathbf{M}^{-p})]^{1/p}$, and thus to the minimization of $\mathrm{tr}(\mathbf{M}^{-p})$ when p is positive. A classical example is A-optimal design, which corresponds to $p = 1$. Taking the limit of $\Phi_p^+(\cdot)$ when p tends to zero, we obtain $\Phi_0^+(\mathbf{M}) = [\det(\mathbf{M})]^{1/m}$, which corresponds to D-optimal design. The limit when p tends to infinity gives $\Phi_\infty(\mathbf{M}) = \lambda_{\min}(\mathbf{M})$, the minimum eigenvalue of $\mathbf{M}$, and corresponds to E-optimal design. Some basic properties of ϕ_p-optimal designs are briefly recalled in Sect. 2.

Classical algorithms for optimal design usually apply to situations where $\mathscr{X}$ is a finite set. The performance of the algorithm (in particular, its execution time for a given required precision on $\phi(\cdot)$) then heavily depends on the number k of elements in $\mathscr{X}$. The case of D-optimal design has retained much attention, see, for instance, Ahipasaoglu et al. (2008), Todd and Yildirim (2007), Yu (2010) and Yu (2011). Harman and Pronzato (2007) show how any nonsingular design on $\mathscr{X}$ yields a simple inequality that must be satisfied by the support points of a D-optimal design ξ_0^*. Whatever the iterative method used for the construction of ξ_0^*, this delimitation of the support of ξ_0^* permits to reduce the cardinality of $\mathscr{X}$ along the iterations, with the inequality becoming more stringent when approaching the optimum, hence producing a significant acceleration of the algorithm. Put in other words, the delimitation of the support of an optimal design facilitates the optimization by focussing the search on the useful part of the design space $\mathscr{X}$. The objective of the paper is to extend the results in Harman and Pronzato (2007) to the ϕ_p-class (1) of design criteria. The condition obtained does not tell what the optimum support is, but indicates where it cannot be.

The paper is organized as follows. Section 2 recalls the main properties of ϕ_p-optimal design that are useful for the rest of the paper. The main result is derived in Sect. 3 and illustrative examples are given in Sect. 4. Finally, Sect. 5 concludes and indicates some possible extensions. The technical parts of the proofs are given in appendix.

2 Some basic properties of ϕ_p-optimal designs

The criteria $\Phi_p^+(\cdot)$ defined by (1) satisfy $\Phi_p^+(\mathbf{I}_m) = 1$ for $\mathbf{I}_m$ the m-dimensional identity matrix and $\Phi_p^+(a\mathbf{M}) = a\,\Phi_p^+(\mathbf{M})$ for any $a > 0$ and any $\mathbf{M} \in \mathbb{M}$. Note that, from Caratheodory's theorem, a finitely-supported optimal design always exists, with $m(m+1)/2$ support points at most. We also have the following properties.

Lemma 1 *For any $p \in (-1, \infty)$, the criterion $\Phi_p^+(\cdot)$ satisfies the following:*

(i) $\Phi_p^+(\cdot)$ is strictly concave on the set $\mathbb{M}^+$ of symmetric positive definite $m \times m$ matrices; it is strictly isotonic on $\mathbb{M}$ for $p \in (-1, 0)$ and strictly isotonic on $\mathbb{M}^+$ for $p \in [0, \infty)$.

(ii) Any ϕ_p-optimal design ξ_p^* is nonsingular.

(iii) The optimal matrix $\mathbf{M}_* = \mathbf{M}_*[p]$ is unique.

Part (i) is proved in Pukelsheim (1993, Chap. 6) and the proof of (ii) is given in Appendix; (iii) is a direct consequence of (i) and (ii): since an optimal design matrix $\mathbf{M}_*$ is nonsingular, the strict concavity of $\Phi_p^+(\cdot)$ at $\mathbf{M}_*$ implies that $\mathbf{M}_*$ is unique. Note that this does not imply that the optimal design measure ξ_p^* maximizing $\phi_p(\xi)$ is unique.

We shall only consider values of p in $(-1, \infty)$ and, from Lemma 1-(ii), we can thus restrict our attention to matrices $\mathbf{M}$ in $\mathbb{M}^+$. $\Phi_p^+(\cdot)$ is differentiable at any $\mathbf{M} \in \mathbb{M}^+$, with gradient

$$\nabla\Phi_p^+(\mathbf{M}) = \frac{1}{m}\,[\Phi_p^+(\mathbf{M})]^{p+1}\,\mathbf{M}^{-(p+1)} = \frac{\Phi_p^+(\mathbf{M})}{\mathrm{tr}(\mathbf{M}^{-p})}\,\mathbf{M}^{-(p+1)}\,.$$

The directional derivative $F_{\phi_p}(\xi;\nu) = \lim_{\alpha\to 0^+}(1/\alpha)\{\phi_p[(1-\alpha)\xi + \alpha\nu] - \phi_p(\xi)\}$ is well defined and finite for any $\xi \in \Xi^+$ and any $\nu \in \Xi$, with

$$F_{\phi_p}(\xi;\nu) = \mathrm{tr}\{[\mathbf{M}(\nu) - \mathbf{M}(\xi)]\nabla\Phi_p^+[\mathbf{M}(\xi)]\} = \phi_p(\xi)\,\left\{\frac{\int_{\mathscr{X}} \mathbf{x}^\top\mathbf{M}^{-(p+1)}(\xi)\mathbf{x}\,\nu(\mathrm{d}\mathbf{x})}{\mathrm{tr}[\mathbf{M}^{-p}(\xi)]} - 1\right\}.$$

We shall denote by $F_{\phi_p}(\xi,\mathbf{x}) = F_{\phi_p}(\xi;\delta_\mathbf{x})$ the directional derivative of $\phi_p(\cdot)$ at ξ in the direction of the delta measure at $\mathbf{x}$,

$$F_{\phi_p}(\xi,\mathbf{x}) = \phi_p(\xi)\,\left\{\frac{\mathbf{x}^\top\mathbf{M}^{-(p+1)}(\xi)\mathbf{x}}{\mathrm{tr}[\mathbf{M}^{-p}(\xi)]} - 1\right\}. \tag{2}$$

The following theorem, which relies on the concavity and differentiability of $\Phi_p^+(\cdot)$, is a classical result in optimal design theory, see, e.g., Kiefer (1974) and Pukelsheim (1993, Chap. 7).

Theorem 1 (Equivalence Theorem) *For any $p \in (-1, \infty)$, the following statements are equivalent:*

(i) ξ_p^* *is* ϕ_p*-optimal.*

(ii) $\mathbf{x}^\top \mathbf{M}^{-(p+1)}(\xi_p^*)\mathbf{x} \leq \mathrm{tr}[\mathbf{M}^{-p}(\xi_p^*)]$ *for all* $\mathbf{x} \in \mathscr{X}$*.*

(iii) ξ_p^* *minimizes* $\max_{\mathbf{x}\in\mathscr{X}} F_{\phi_p}(\xi, \mathbf{x})$ *with respect to* $\xi \in \Xi^+$*.*

3 A necessary condition for support points of ϕ_p-optimal designs

3.1 A lower bound on $\mathbf{x}^\top \mathbf{M}^{-(p+1)}\mathbf{x}$ for the support points of an optimal design

Take any $p \in (-1, \infty)$ and any $\xi \in \Xi^+$. We shall omit the dependence in ξ when there is no ambiguity and simply write $\mathbf{M} = \mathbf{M}(\xi)$, $\phi_p = \phi_p(\xi)$. We shall also denote

$$t = t(\xi, p) = \mathrm{tr}[\mathbf{M}^{-p}]\,, \quad t_* = t_*(p) = \mathrm{tr}(\mathbf{M}_*^{-p})\,,$$

with $\mathbf{M}_*$ the optimal matrix satisfying $\phi_p^* = \Phi_p^+(\mathbf{M}_*) = \max_{\nu\in\Xi} \Phi_p^+[\mathbf{M}(\nu)]$. Define

$$\epsilon = \epsilon(\xi, p) = \max_{\mathbf{x}\in\mathscr{X}} \{\mathbf{x}^\top \mathbf{M}^{-(p+1)}\mathbf{x}\} - t\,. \tag{3}$$

The concavity of $\Phi_p^+(\cdot)$ implies that $\phi_p \leq \phi_p^* \leq \phi_p + F_{\phi_p}(\xi; \xi_p^*) \leq \phi_p(\xi) + \max_{x\in\mathscr{X}} F_{\phi_p}(\xi, \mathbf{x})$, with ξ_p^* denoting a ϕ_p-optimal design measure; that is,

$$\phi_p \leq \phi_p^* \leq \phi_p\,(1 + \epsilon/t)\,, \tag{4}$$

see (2).

Since $\mathbf{x}^\top \mathbf{M}^{-(p+1)}\mathbf{x} \leq t + \epsilon$ for all $\mathbf{x} \in \mathscr{X}$, see (3), we have

$$\mathrm{tr}[\mathbf{M}_*\mathbf{M}^{-(p+1)}] \leq t + \epsilon\,. \tag{5}$$

On the other hand, the optimality of ξ_p^* implies (see Th. 1-(*ii*))

$$\mathrm{tr}[\mathbf{M}\mathbf{M}_*^{-(p+1)}] \leq t_*\,. \tag{6}$$

Moreover, any support point $\mathbf{x}_*$ of ξ_p^* satisfies $\mathbf{x}_*^\top \mathbf{M}_*^{-(p+1)}\mathbf{x}_* = t_*$. We use a construction similar to that in Harman and Pronzato (2007) and define $\mathbf{H} = \mathbf{H}(\xi, p) = \mathbf{M}^{-(p+1)/2}\mathbf{M}_*^{p+1}\mathbf{M}^{-(p+1)/2}$. Then we can write

$$\mathbf{x}_*^\top \mathbf{M}^{-(p+1)}\mathbf{x}_* = \mathbf{x}_*^\top \mathbf{M}^{-(p+1)/2}\mathbf{H}^{-1/2}\mathbf{H}\mathbf{H}^{-1/2}\mathbf{M}^{-(p+1)/2}\mathbf{x}_* \geq \lambda_1\, \mathbf{x}_*^\top \mathbf{M}_*^{-(p+1)}\mathbf{x}_* = \lambda_1\, t_*\,,$$

with $\lambda_1 = \lambda_1(\xi, \xi_p^*, p) = \lambda_{\min}(\mathbf{H})$, the minimum eigenvalue of $\mathbf{H}$. Notice that $\lambda_1 > 0$. λ_1 depends on $\mathbf{M}_*$ which is unknown. Below we shall construct a lower bound $\underline{\lambda_1}$ on λ_1 and thus obtain a necessary condition for support points $\mathbf{x}_*$ of ξ_p^*, in the form:

$$\mathbf{x}_*^\top \mathbf{M}^{-(p+1)}\mathbf{x}_* \geq \underline{\lambda_1}\, t_*\,. \tag{7}$$

When $p = 0$ (D-optimal design), we have $t = t_* = m$, and this necessary condition is simply

$$\mathbf{x}_*^\top \mathbf{M}^{-1} \mathbf{x}_* \geq \underline{\lambda_1}\, m \quad (p = 0)\,; \tag{8}$$

it corresponds to the case treated in Harman and Pronzato (2007). When $p \neq 0$, t_* is usually unknown and we shall use

$$\begin{aligned} \mathbf{x}_*^\top \mathbf{M}^{-(p+1)} \mathbf{x}_* &\geq \underline{\lambda_1}\, t(1+\epsilon/t)^{-p} \quad \text{for } p > 0\,, &(9)\\ \mathbf{x}_*^\top \mathbf{M}^{-(p+1)} \mathbf{x}_* &\geq \underline{\lambda_1}\, t \quad\quad\quad\quad\quad \text{for } -1 < p < 0\,, &(10) \end{aligned}$$

see (4) and the definitions of t, t_*, ϕ_p, ϕ_p^*. Next section is devoted to the construction of the lower bound $\underline{\lambda_1}$, using the inequalities (5) and (6).

3.2 Construction of the lower bound $\underline{\lambda_1}$

The inequality (5) can be rewritten as $\mathrm{tr}(\mathbf{H}^{1/(p+1)}\mathbf{M}^{-p}) \leq t + \epsilon$ and (6) can be rewritten as $\mathrm{tr}(\mathbf{H}^{-1}\mathbf{M}^{-p}) \leq t_*$. Consider the spectral decomposition $\mathbf{H} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^\top$, with $\mathbf{S}\mathbf{S}^\top = \mathbf{S}^\top\mathbf{S} = \mathbf{I}_m$ and $\mathbf{\Lambda}$ the diagonal matrix whose diagonal elements are the eigenvalues λ_i of $\mathbf{H}$ sorted by increasing values. Denote $\mathbf{B} = \mathbf{S}^\top \mathbf{M}^{-p} \mathbf{S}$ and $b_i = \{\mathbf{B}\}_{ii}$ its diagonal elements, $i = 1, \ldots, m$. $\mathbf{B}$ has the same set of eigenvalues as $\mathbf{M}^{-p}$ and

$$0 < \underline{b_1} = \lambda_{\min}(\mathbf{M}^{-p}) \leq b_i \leq \lambda_{\max}(\mathbf{M}^{-p})\,,\ i = 1, \ldots, m\,, \tag{11}$$

see, e.g., Magnus and Neudecker (1999, p. 211). We then obtain that (5) and (6) are respectively equivalent to

$$\begin{aligned} \sum_{i=1}^m b_i\, \lambda_i^{1/(p+1)} &\leq t + \epsilon\,,\\ \sum_{i=1}^m b_i/\lambda_i &\leq t_*\,. \end{aligned} \tag{12}$$

Remark 1 *Inequality* (12) *implies that* $\underline{\lambda_1} \geq b_1/t_* \geq \underline{b_1}/t_*$. *When plugged in* (7)*, it gives* $\mathbf{x}_*^\top \mathbf{M}^{-(p+1)} \mathbf{x}_* \geq \underline{b_1}$. *Although this bound is rather loose for* $m \geq 2$*, it cannot be improved when* $m = 1$. *Indeed,* $m = 1$ *implies* $\underline{b_1} = b_1 = t$ *and the inequality* $\mathbf{x}_*^\top \mathbf{M}^{-(q+1)} \mathbf{x}_* \geq t$ *is the tightest we can obtain, see Th. 1-(ii). In the following we shall suppose that* $m \geq 2$.

Denote $\omega_i = \lambda_i^{1/(p+1)}$ for $i = 1, \ldots, m \geq 2$. The determination of $\underline{\lambda_1}$ amounts to the solution of the following optimization problem: minimize ω_1 with respect to $\omega = (\omega_1, \ldots, \omega_m)^\top$ under the constraints $0 \leq \omega_1 \leq \omega_2 \leq \cdots \leq \omega_m$, $\sum_{i=1}^m b_i\, \omega_i \leq t + \epsilon$ and $\sum_{i=1}^m b_i/\omega_i^{p+1} \leq t_*$. This is a convex problem, with Lagrangian

$$L(\omega, \mu_1, \mu_2) = \omega_1 + \mu_1 \left(\sum_{i=1}^m b_i\, \omega_i - t - \epsilon \right) + \mu_2 \left(\sum_{i=1}^m b_i/\omega_i^{p+1} - t_* \right)\,.$$

Its stationarity with respect to ω indicates that the optimal solution satisfies $\omega_i = \omega_2$ for all $i \geq 2$. Since $\sum_{i=1}^m b_i = \mathrm{tr}(\mathbf{M}^{-p}) = t$, from the Kuhn-Tucker conditions we obtain

$$\begin{aligned} b_1\,\omega_1 + (t-b_1)\omega_2 &= t+\epsilon\,, \\ b_1/\omega_1^{p+1} + (t-b_1)/\omega_2^{p+1} &= t_*\,, \end{aligned}$$

or equivalently

$$\alpha\,\omega_1 + (1-\alpha)\omega_2 = 1+\beta\,, \tag{13}$$

$$\alpha/\omega_1^{p+1} + (1-\alpha)/\omega_2^{p+1} = \gamma_*\,, \tag{14}$$

where $\alpha = b_1/t$, $\beta = \epsilon/t \geq 0$ and $\gamma_* = t^*/t$.

When $p = 0$ (D-optimal design), then $\alpha = 1/m$, $\gamma_* = 1$ and (13), (14) can be directly solved for ω_1, ω_2, yielding $\underline{\lambda_1} = \omega_1$ to be used in (8), see Harman and Pronzato (2007). However, when $p \neq 0$, α depends on $\mathbf{M}_*$ and γ_* depends on t_* and are thus usually unknown. We must then determine the lowest value of $\omega_1 \leq \omega_2$ satisfying (13), (14) given the information available on α and γ_*; that is, respectively, (11) which gives $1 > \alpha \geq \underline{b_1}/t = \lambda_{\min}(\mathbf{M}^{-p})/\mathrm{tr}(\mathbf{M}^{-p})$, and (4) which implies that γ_* satisfies

$$\gamma_* \in [(1+\beta)^{-p}, 1] \qquad \text{if } p \geq 0\,, \tag{15}$$

$$\gamma_* \in [1, (1+\beta)^{-p}] \qquad \text{if } p \leq 0\,. \tag{16}$$

The solution to this problem is given in appendix and yields the main result of the paper.

Theorem 2 *For any $p \in (-1, \infty)$ and any design $\xi \in \Xi^+$, any point $\mathbf{x}_* \in \mathscr{X}$ such that*

$$\mathbf{x}_*^\top \mathbf{M}^{-(p+1)}(\xi)\mathbf{x}_* < C(\xi, p) = \omega_1^{p+1}\, B(t, \epsilon) \tag{17}$$

cannot be support point of a ϕ_p-optimal design measure ξ_p^*, where we denoted $t = \mathrm{tr}[\mathbf{M}^{-p}(\xi)]$, $\epsilon = \max_{\mathbf{x}\in\mathscr{X}} \mathbf{x}^\top \mathbf{M}^{-(p+1)}(\xi)\mathbf{x} - t$, $B(t,\epsilon) = t\,\min\{1, (1+\epsilon/t)^{-p}\}$, and where ω_1 is the unique solution for θ in the interval $((\alpha/\gamma)^{1/(p+1)}, (1/\gamma)^{1/(p+1)}]$ of the equation

$$F(\theta; \alpha, \epsilon, t, \gamma, p) = \frac{\alpha}{\theta^{p+1}} + \frac{(1-\alpha)^{p+2}}{(1+\epsilon/t - \alpha\theta)^{p+1}} - \gamma = 0 \tag{18}$$

with $\alpha = \lambda_{\min}[\mathbf{M}^{-p}(\xi)]/\mathrm{tr}[\mathbf{M}^{-p}(\xi)]$ and $\gamma = \max\{1, (1+\epsilon/t)^{-p}\}$.

In the special case when $t_* = \mathrm{tr}[\mathbf{M}^{-p}(\xi_p^*)]$ is known (thus in particular if $p = 0$), one can take $B(t,\epsilon) = t_*$ and $\gamma = \gamma_* = t^*/t$ in (17), (18).

Remark 2

1. *When p is integer, $F(\theta;\alpha,\epsilon,t,\gamma,p)=0$ is a polynomial equation in θ of degree $2(p+1)$.*

2. *From the definition* (3) *of $\epsilon=\epsilon(\xi,p)$, $\delta=\max_{\mathbf{x}\in\mathscr{X}} F_{\phi_p}(\xi,\mathbf{x})=\epsilon\,\phi_p(\xi)/t$, see* (2)*, and* (17) *is equivalent to $F_{\phi_p}(\xi,\mathbf{x}_*)<h_p[\mathbf{M}(\xi),\delta]=\phi_p(\xi)\,[C(\xi,p)/t-1]$. Note that $C(\xi,p)\leq t$, so that all points $\mathbf{x}$ such that $F_{\phi_p}(\xi,\mathbf{x})\geq 0$ are potential support points of ξ_p^*.*

3. *Suppose $p>0$ with t_* unknown and $\epsilon\to\infty$; then, $B(t,\epsilon)\to 0$, so that $C(\xi,p)\to 0$ and the condition* (17) *brings no information on the support of ξ_p^*. The same is true when $p<0$ with t_* unknown and $\epsilon\to\infty$: $\gamma\to\infty$, so that $\omega_1\to 0$ and again $C(\xi,p)\to 0$. Suppose now that t_* is known. Then, $C(\xi,p)=t_*\,\omega_1^{p+1}\in(\lambda_{\min}[\mathbf{M}^{-p}(\xi)],\,\mathrm{tr}[\mathbf{M}^{-p}(\xi)]]$ and $\omega_1^{p+1}\to\alpha/\gamma_*=\lambda_{\min}[\mathbf{M}^{-p}(\xi)]/t_*$ as $\epsilon\to\infty$, see* (18)*, so that $C(\xi,p)\to\underline{b_1}=\lambda_{\min}[\mathbf{M}^{-p}(\xi)]$ and we recover the same bound as in Remark 1.*

4. *Using a construction similar to that in Harman and Pronzato (2007, Th. 3), one can show that the bound* (17) *with $B(t,\epsilon)=t_*$ and $\gamma=t^*/t$ gives the tightest necessary condition for support points: for any $m\geq 2$, any $\epsilon>0$ and any $\delta>0$, one can exhibit an example with a design space $\mathscr{X}$, a design measure ξ such that $\max_{\mathbf{x}\in\mathscr{X}}\{\mathbf{x}^\top\mathbf{M}^{-(p+1)}\mathbf{x}\}-t=\epsilon$, and an optimal design ξ_p^* with support point $\mathbf{x}_*$ such that $\mathbf{x}_*^\top\mathbf{M}^{-(p+1)}\mathbf{x}_*<\omega_1^{p+1}\,t_*+\delta$ (with $\mathbf{M}$ and $\mathbf{M}_*$ diagonal and $\mathbf{H}$ having eigenvalues $\lambda_1<\lambda_2=\cdots=\lambda_m$).*

4 Examples

Example 1. Consider the linear regression model with $\mathbf{x}=\mathbf{x}(s)=(1,\ s,\ s^2)^\top$, $s\in[-1,1]$ ($m=3$). For any $p\in(-1,\infty)$, the ϕ_p-optimal design on $[-1,1]$ is unique and is supported at the three points $\{-1,0,1\}$. For symmetry reasons, it corresponds to

$$\xi_\tau=\tau\,\delta_{-1}+(1-2\tau)\,\delta_0+\tau\,\delta_1$$

for some particular $\tau^*=\tau^*(p)$, with $\tau^*(-1/2)=0.45$, $\tau^*(0)=1/3$ (D-optimal design), $\tau^*(1)=1/4$ (A-optimal design) and, in the limit $p\to\infty$, $\tau^*(\infty)=0.2$ (E-optimal design), see Fig. 2-left for a plot of $\tau^*(p)$ for $p\in[-1/2,1]$. Here, δ_s denotes the Dirac delta measure at s.

To illustrate the impact of not knowing t_* on the construction of ω_1^{p+1} through the solution of (18), we take $p=1$ and compute ω_1^{p+1} for the cases $\gamma=1$ (t_* unknown) and $\gamma=t^*/t$ (t_* known) for different designs ξ_τ, $\tau\in[\tau^*(1)-1/16,\,\tau^*(1)+1/16]$. Figure 1 shows that the value obtained for t_* unknown (solid line) is not much worse, i.e., smaller, than the value for t_* known (dashed line). Note that considering different designs ξ_τ with $\tau\neq\tau^*(p)$ is equivalent to considering different ϵ given by (3), with ϵ being approximately linear in $|\tau-\tau^*(p)|$ for the range of values of τ considered.

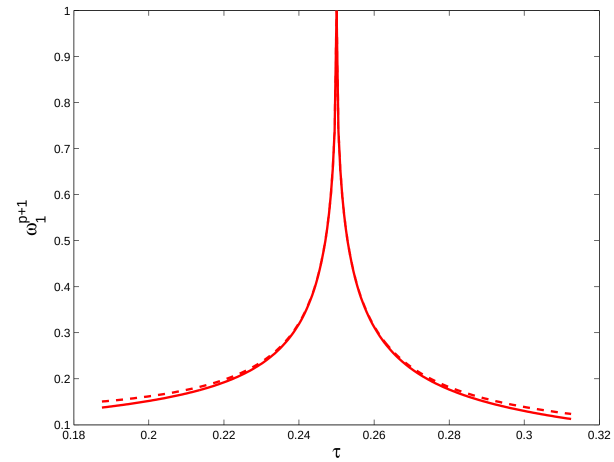

Figure 1: Value of ω_1^{p+1} for different designs ξ_τ, $\tau \in [3/16,\ 5/16]$ ($p = 1$, t_* unknown in *solid line*, t_* known in *dashed line*).

The marginal deterioration of the bound (17) due to the ignorance of t_* when ϵ is small enough is further illustrated by Fig. 2. Here, we set ϵ at some fixed value (the values $\epsilon = 0.1$ and $\epsilon = 1$ are considered), and for values of p in the range $[-1/2, 1]$ we compute $\tau(p, \epsilon)$ such that $\max_{s\in[-1,1]} \mathbf{x}^\top(s)\mathbf{M}^{-(p+1)}(\xi_\tau)\mathbf{x}(s) = \mathrm{tr}[\mathbf{M}^{-p}(\xi_\tau)] + \epsilon$. The values of $\tau^*(p)$ and $\tau(p, 0.1)$ are shown in Fig. 2-left, in solid and dashed lines respectively. Then, for each p and associated design $\xi_{\tau(p,\epsilon)}$ we compute the bound $C(\xi_{\tau(p,\epsilon)}, p)$ of (17) in the two situations t_* unknown and t_* known; see the plots in Fig. 2-right. Note that bound for t_* unknown (solid line) remains near the bound for t_* known (dashed line) when $\epsilon = 0.1$; the situation deteriorates for larger ϵ but the two bounds get close as p approaches 0 and exactly coincide at $p = 0$ (since then $t = t_* = m$).

Example 2. Take now the complete product-type interaction model $\mathbf{x}(\mathbf{s}) = \mathbf{x}(s_1) \otimes \mathbf{x}(s_2)$, $\mathbf{s} = (s_1,\ s_2)$, with $\otimes$ denoting tensor product and $\mathbf{x}(s_i) = (1,\ s_i,\ s_i^2)^\top$, $s_i \in [-1, 1]$, for $i = 1, 2$ ($m = 9$). The D-optimal (respectively A-optimal) design for this problem is the cross product of two D-optimal designs (resp. A-optimal designs) for one single factor, i.e., it corresponds to the cross product of two designs ξ_τ with $\tau = 1/3$ (resp. $\tau = 1/4$), see Schwabe (1996, Chap. 4 and 5). The optimal values of ϕ_p, $p = 0, 1$, are $\phi_0^* = 16^{1/3}/9 \simeq 0.2800$ and $\phi_1^* = 9/64 \simeq 0.1406$.

We consider the iterative construction of optimal designs through the recursion

$$w_i^{k+1} = w_i^k \frac{[\mathbf{x}_i^\top \mathbf{M}^{-(p+1)}(\xi_k)\mathbf{x}_i]^a}{\sum_{i=1}^{N_k} [\mathbf{x}_i^\top \mathbf{M}^{-(p+1)}(\xi_k)\mathbf{x}_i]^a}\,, \qquad (19)$$

where ξ_k, the design measure at iteration k, allocates mass w_i^k at the point $\mathbf{x}_i$ present in $\mathscr{X}$ at iteration k, $i = 1, \ldots, N_k$. The initial design space corresponds to a uniform grid for $\mathbf{s}$, with s_i varying from -1 to 1 by steps of 0.05 (41 values), $i = 1, 2$, which gives $N_0 = 1681$. The initial

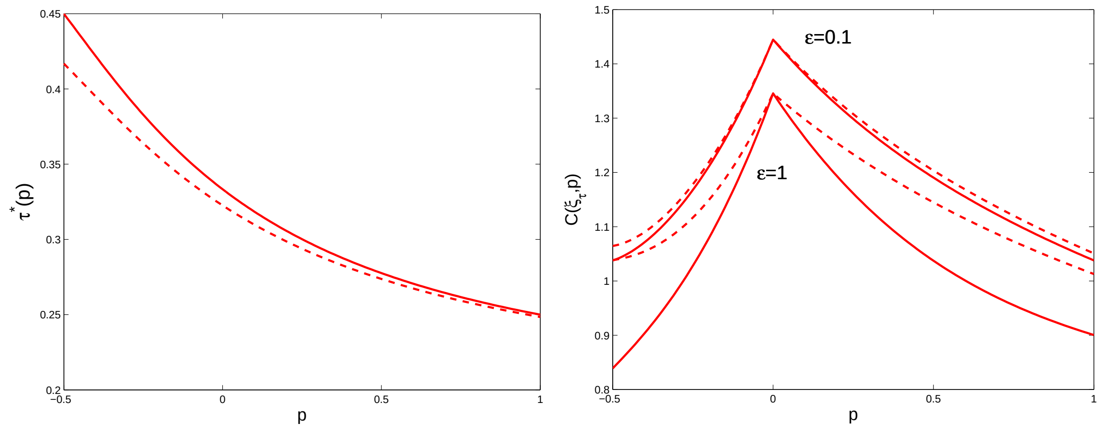

Figure 2: Left: $\tau^*(p)$ such that $\xi_{\tau^*(p)} = \xi_p^*$ is ϕ_p-optimal for p (*solid line*) and $\tau(p,\epsilon)$ such that $\max_{s\in[-1,1]} \mathbf{x}^\top(s)\mathbf{M}^{-(p+1)}(\xi_{\tau(p,\epsilon)})\mathbf{x}(s) = \mathrm{tr}[\mathbf{M}^{-p}(\xi_{\tau(p,\epsilon)})] + \epsilon$ ($\epsilon = 0.1$, *dashed line*). Right: bound $C(\xi_{\tau(p,\epsilon)}, p)$ in (17) for the two cases t_* unknown (*solid line*) and t_* known (*dashed line*) for $\epsilon = 0.1$ and $\epsilon = 1$.

design ξ_0 is the uniform measure on those 1681 points. We take $a = 1$ for D-optimal design ($p = 0$) and $a = 1/2$ for A-optimal design ($p = 1$), which ensures monotonic convergence to the optimum, see Titterington (1976) and Pázman (1986) for D-optimal design and Torsney (1983) for A-optimal design; see also Fig. 3-left. Due to the convergence of ξ_k to the optimal design, $\epsilon_k = \epsilon(\xi_k)$ given by (3) is decreasing with k, see Fig. 3-right.

We use inequality (17) to reduce the cardinality N_k of $\mathscr{X}$ when possible: any point that violates (17) cannot be a support point of the optimal measure and is removed from $\mathscr{X}$. Here we simply set its mass to zero and rescale the weights of remaining point so that they sum to one, but more sophisticated reallocation rules can be used, see Harman and Pronzato (2007). N_k thus decreases with k, rendering the iterations (19) simpler and simpler as k increases. Figure 4 shows the evolution of N_k with k, both for D-optimal and A-optimal designs. The decrease of N_k is slower for the latter, the bound $C(\xi, p)$ in (17) being more pessimistic, see Fig. 2-right, and ϵ being larger, see Fig.3-right. Note that the cancelation of points does not hamper the convergence of (19) since (17) is used a finite number of times only (obviously bounded by N_0) — the heuristic rule used to reallocate weights of points that are removed may, however, impact monotonicity, although this is not the case in the present example, see Fig. 3-left.

5 Possible extensions and conclusions

Multivariate regression and Bayesian optimal design involve information matrices that can be expressed as $\mathbf{M}(\xi) = \int_{\mathscr{X}} \mathscr{M}(\mathbf{x})\,\xi(\mathrm{d}\mathbf{x})$ with $\mathscr{M}(\mathbf{x}) \in \mathbb{M}$ having rank larger than one (we suppose that $\mathscr{M}(\cdot)$ is measurable and that $\{\mathscr{M}(\mathbf{x}), \mathbf{x} \in \mathscr{X}\}$ forms a compact subset of $\mathbb{M}$). The

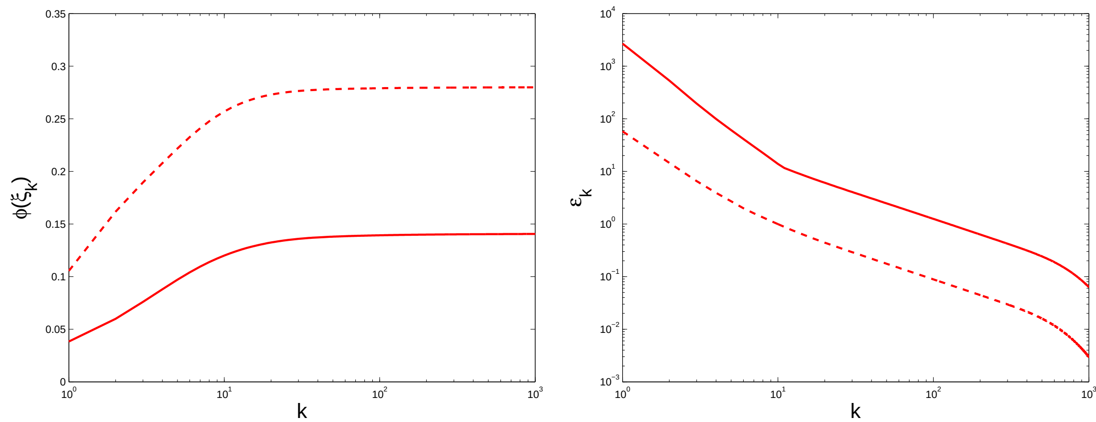

Figure 3: $\phi(\xi_k)$ — left — and $\epsilon_k = \epsilon(\xi_k)$ given by (3) — right — as functions of k for the recursion (19); D-optimal design is in *dashed line*, A-optimal design is in *solid line*.

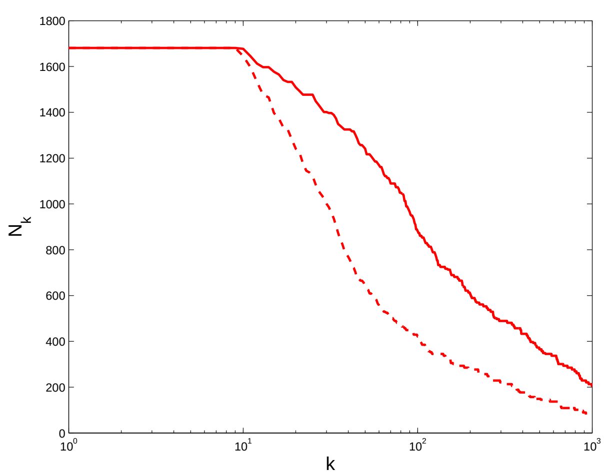

Figure 4: N_k as a function of k when using (17) to remove points from $\mathscr{X}$; for D-optimal design (*dashed line*) and A-optimal design (*solid line*).

results presented here can easily be extended to that situation, following the same lines as in Harman and Trnovská (2009) where the case $p = 0$ is considered.

The E-optimality criterion $\phi_E(\xi) = \Phi_E[\mathbf{M}(\xi)] = \lambda_{\min}[\mathbf{M}(\xi)]$ is not differentiable in general, but $\Phi_E(\cdot)$ is differentiable at $\mathbf{M}$ when $\lambda_{\min}(\mathbf{M})$ has multiplicity one, with gradient $\nabla\phi_E(\mathbf{M}) = \mathbf{v}\mathbf{v}^\top$ where $\mathbf{v}$ denotes the eigenvector of unit length (unique up to a sign change) associated with $\lambda_{\min}(\mathbf{M})$. Although $\phi_E(\xi)$ corresponds to the limit of $\phi_p^+(\xi)$ as p tends to infinity, the results of Sect. 3 do not extend to this limiting situation, even in the differentiable case; E-optimality thus requires a special treatment and will be considered elsewhere.

The determination of a D-optimal design can be used for maximum-likelihood estimation in mixture models, see, e.g., Lindsay (1983) and Mallet (1986), and for the construction of the minimum-volume ellipsoid containing a compact set, see, e.g., Sibson (1972), Khachiyan and Todd (1993) and Khachiyan (1996). More generally, for any $q \in (-1, \infty)$ the determination of the ellipsoid $\mathscr{E}(\mathbf{A}) = \{\mathbf{z} \in \mathbb{R}^m : \mathbf{z}^\top\mathbf{A}\mathbf{z} \leq 1\}$, $\mathbf{A} \in \mathbb{M}$, containing the k points $\mathbf{x}_1, \ldots, \mathbf{x}_k$ of $\mathbb{R}^m$ and such that $\phi_q(\mathbf{A})$ is maximum is equivalent to the determination of a ϕ_p-optimal design on $\mathscr{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ with $p = -q/(1+q) \in (-1, \infty)$, and the optimal matrix $\mathbf{A}_*$ equals $\mathbf{M}_*^{-(p+1)}/t_*$; see Pukelsheim (1993, Chap. 6). The delimitation of the support points of a ϕ_p-optimal design can therefore also be used to accelerate the algorithmic construction of "ϕ_q-optimal ellipsoids" containing compact sets.

In Sect. 4, we considered the suppression of points that cannot be support points of an optimal design in a multiplicative algorithm. When $\mathscr{X}$ is not finite, or is finite but very large, it is advisable to use a vertex-direction or a vertex-exchange algorithm, see, e.g., Fedorov (1972), Wu (1978) and Böhning (1986). This requires the determination at each iteration, say iteration k, of a point $\hat{\mathbf{x}}_k$ of $\mathscr{X}$ that maximizes $F_{\phi_p}(\xi_k, \mathbf{x})$ given by (2), at least approximately. Condition (17) of Theorem 2 can then be used to restrict the search for a suitable $\hat{\mathbf{x}}_k$ in a domain that shrinks as k increases. Further developments are required to construct algorithms making an efficient use of (17) for the inclusion of new support points.

Appendix

Proof of Lemma 1-(ii).

For $p \geq 0$ the result follows from the observation that $\Phi_p^+(\mathbf{M}) = 0$ when $\mathbf{M}$ is singular while there exists a nonsingular $\mathbf{M}(\xi)$ with $\Phi_p^+[\mathbf{M}(\xi)] > 0$.

For $p \in (-1, 0)$ we prove the result by contradiction. Take any $\mathbf{M}_0 = \mathbf{M}(\xi_0)$ singular and consider its spectral decomposition in an orthonormal basis of eigenvectors $\mathbf{v}_i$, $i = 1, \ldots, m$: $\mathbf{M}_0 = \sum_{i=1}^m \lambda_i \mathbf{v}_i\mathbf{v}_i^\top$. Suppose that the eigenvalues λ_i are sorted by increasing values, so that $\lambda_i = 0$ for $i = 1, \ldots, s$ when the eigenvalue 0 has multiplicity s (and $\mathbf{M}_0$ has rank $m - s$). Since there exists a nonsingular design, $\mathscr{X}$ spans $\mathbb{R}^m$ and each eigenvector $\mathbf{v}_i$, $i = 1, \ldots, s$, can be written as $\mathbf{v}_i = \gamma_i \int_{\mathscr{X} \cup (-\mathscr{X})} \mathbf{z}\, \mu_i(\mathrm{d}\mathbf{z})$ for some $\gamma_i > 0$, where $\mu_i(\cdot)$ is a probability measure on the compact set $\mathscr{X} \cup (-\mathscr{X})$ (in fact, from Caratheodory's theorem, one may consider finitely supported measures only, with $m+1$ support points at most). Then,

$$\gamma_i^2 \int_{\mathscr{X} \cup (-\mathscr{X})} \mathbf{z}\mathbf{z}^\top \, \mu_i(\mathrm{d}\mathbf{z}) - \mathbf{v}_i \mathbf{v}_i^\top = \int_{\mathscr{X} \cup (-\mathscr{X})} (\gamma_i \, \mathbf{z} - \mathbf{v}_i)(\gamma_i \, \mathbf{z} - \mathbf{v}_i)^\top \, \mu_i(\mathrm{d}\mathbf{z}) \,,$$

which is non-negative definite. Denote $\mu(\mathrm{d}\mathbf{z}) = [\sum_{i=1}^s \gamma_i^2 \, \mu_i(\mathrm{d}\mathbf{z})]/(\sum_{i=1}^s \gamma_i^2)$, which defines a probability measure on $\mathscr{X} \cup (-\mathscr{X})$. We thus obtain that, for any $\alpha \in (0,1)$, the matrix

$$\left[(1-\alpha)\, \mathbf{M}_0 + \alpha \int_{\mathscr{X} \cup (-\mathscr{X})} \mathbf{z}\mathbf{z}^\top \, \mu(\mathrm{d}\mathbf{z})\right] - \left[(1-\alpha)\, \mathbf{M}_0 + \alpha \, \frac{\sum_{i=1}^s \mathbf{v}_i \mathbf{v}_i^\top}{\sum_{i=1}^s \gamma_i^2}\right]$$

is non-negative definite. Now, $\int_{\mathscr{X} \cup (-\mathscr{X})} \mathbf{z}\mathbf{z}^\top \, \mu(\mathrm{d}\mathbf{z})$ can be written as $\int_{\mathscr{X}} \mathbf{x}\mathbf{x}^\top \, \tilde\mu(\mathrm{d}\mathbf{x})$, where $\tilde\mu(\mathscr{A}) = \mu(\mathscr{A}) + \mu(-\mathscr{A})$ for any measurable set $\mathscr{A} \subset \mathscr{X}$, and $\tilde\mu(\cdot)$ is thus a design measure on $\mathscr{X}$. Therefore, $\Phi_p^+(\mathbf{M}_\alpha) \geq \Phi_p^+(\mathbf{M}'_\alpha)$, where $\mathbf{M}_\alpha = \mathbf{M}[(1-\alpha)\xi_0 + \alpha\tilde\mu] = (1-\alpha)\, \mathbf{M}_0 + \alpha \int_{\mathscr{X}} \mathbf{x}\mathbf{x}^\top \, \tilde\mu(\mathrm{d}\mathbf{x})$ and $\mathbf{M}'_\alpha = (1-\alpha)\, \mathbf{M}_0 + (\alpha/\rho) \sum_{i=1}^s \mathbf{v}_i \mathbf{v}_i^\top$, with $\rho = \sum_{i=1}^s \gamma_i^2$. The eigenvector decomposition of $\mathbf{M}'_\alpha$ gives

$$\Phi_p^+(\mathbf{M}'_\alpha) = \left\{ \frac{1}{m} \left[(1-\alpha)^{-p} \left(\sum_{i=s+1}^m \lambda_i^{-p} \right) + \alpha^{-p} \frac{s}{\rho^{-p}} \right] \right\}^{-1/p}$$

which reaches its maximum value for $\alpha = \alpha^* = [1 + (\rho^{-p} \sum_{i=s+1}^m \lambda_i^{-p}/s)^{1/(p+1)}]^{-1} \in (0,1)$. It implies that $\Phi_p^+(\mathbf{M}_{\alpha^*}) \geq \Phi_p^+(\mathbf{M}'_{\alpha^*}) > \Phi_p^+(\mathbf{M}_0)$; that is, $\phi_p[(1-\alpha^*)\xi_0 + \alpha^*\tilde\mu] > \phi_p(\xi_0)$ and ξ_0 is not optimal. ■

Proof of Th 2. The proof is in three parts. In (i) we show that for given α and γ_* the equations (13), (14) with $\omega_1 \leq \omega_2$ have a unique solution $\omega_1^*(\alpha, \gamma_*)$ for ω_1, with $\omega_1^*(\alpha, \gamma_*) \in ((\alpha/\gamma_*)^{1/(p+1)}, (1/\gamma_*)^{1/(p+1)}]$. Then in (ii) we show that this solution is non-decreasing in α, so that the required lowest bound is obtained for $\alpha = \underline{b_1}/t$, see (11). Finally, in (iii) we consider the case when t_* is unknown.

(i) Expressing ω_2 as a function of ω_1 using (13), we obtain $\omega_2 = f_1(\omega_1) = (1+\beta-\alpha\omega_1)/(1-\alpha)$, i.e., a decreasing linear function of ω_1 with slope $-\alpha/(1-\alpha)$ and such that $f_1[(1+\beta)/\alpha] = 0$. Doing the same with (14), we obtain $\omega_2 = f_2(\omega_1)$ with $f_2(\cdot)$ decreasing and concave for $\omega_1 \in ((\alpha/\gamma_*)^{1/(p+1)}, \infty)$, $f_2(\theta)$ tending to infinity when θ approaches $(\alpha/\gamma_*)^{1/(p+1)}$ from above and $\lim_{\theta\to\infty} f_2(\theta) = 1/\alpha - 1$. Note that (15), (16) imply that $(\alpha/\gamma_*)^{1/(p+1)} < (1/\gamma_*)^{1/(p+1)} < (1+\beta)/\alpha$. Therefore, $f_2(\theta) > f_1(\theta)$ for θ close enough to $(\alpha/\gamma_*)^{1/(p+1)}$ or large enough.

Denote $f_2'(\theta) = \mathrm{d}f_2(\theta)/\mathrm{d}\theta$ and consider $\theta_* = (1/\gamma_*)^{1/(p+1)}$. Direct calculations indicate that $f_2(\theta_*) = \theta_*$, $f_2'(\theta_*) = -\alpha/(1-\alpha)$ with, moreover, $f_1(\theta_*) > f_2(\theta_*)$ when $\beta > 0$, i.e., when $\epsilon > 0$, due to (15) and (16). Two solutions $\omega_{1,a}^*, \omega_{1,b}^*$ thus exist for (13), (14), with $\omega_{1,a}^* < \theta_* < \omega_{1,b}^*$. Only

$\omega_{1,a}^*$ is such that the associated $\omega_{2,a}^*$ satisfies $\omega_{2,a}^* > \omega_{1,a}^*$. When $\epsilon = 0$, then $f_1(\theta_*) = f_2(\theta_*) = \theta_*$ and the two solutions $\omega_{1,a}^*, \omega_{1,b}^*$ are confounded and equal θ_* (and also coincide with $\omega_{2,a}^*$ and $\omega_{2,b}^*$). The equations (13) and (14) with $\omega_1 \leq \omega_2$ thus always have a unique solution $\omega_1^*(\alpha, \gamma_*)$ and this solution belongs to the interval $((\alpha/\gamma_*)^{1/(p+1)}, \theta_*]$.

(*ii*) Applying the implicit function theorem to (13), (14) we obtain that the solution $\omega_1^*(\alpha, \gamma_*)$ satisfies

$$
\begin{aligned}
\frac{\partial \omega_1^*(\alpha, \gamma_*)}{\partial \alpha} &= \frac{(p+1)(\omega_1^*)^{p+2}(\omega_1^* - \omega_2^*) + \omega_1^* \omega_2^* [(\omega_2^*)^{p+1} - (\omega_1^*)^{p+1}]}{\alpha(p+1)[(\omega_2^*)^{p+2} - (\omega_1^*)^{p+2}]} \\
&= \frac{\omega_1^*}{\alpha(p+1)(z^{p+2} - 1)} \left[(p+1)(1-z) + z(z^{p+1} - 1)\right],
\end{aligned}
$$

where $z = \omega_2^*/\omega_1^* \geq 1$. Denote $f(z) = (p+1)(1-z) + z(z^{p+1}-1)$, its derivative is $\mathrm{d}f(z)/\mathrm{d}z = (p+2)(z^{p+1}-1)$ so that $f(z) \geq f(1) = 0$. Since (11) gives $\alpha \geq \underline{b_1}/t$, one has $\omega_1^*(\alpha, \gamma_*) \geq \omega_1^*(\underline{b_1}/t, \gamma_*)$. The substitution of $[\omega_1^*(\underline{b_1}/t, \gamma_*)]^{p+1}$ for $\underline{\lambda_1}$ in (7) concludes the proof for the case when t_* is known.

(*iii*) When t_* is unknown, an upper bound can be substituted for t_* in (12). Using (15), (16), this amounts at replacing γ_* by the upper bound $\gamma = \max\{1, (1 + \epsilon/t)^{-p}\}$. The necessary conditions (9), (10) with $\underline{\lambda_1} = [\omega_1^*(\underline{b_1}/t, \gamma)]^{p+1}$ then give (17).

References

Ahipasaoglu, S., Sun, P., Todd, M., 2008. Linear convergence of a modified Frank-Wolfe algorithm for computing minimum volume enclosing ellipsoids. Optimization Mehods and Software 23, 5–19.

Böhning, D., 1986. A vertex-exchange-method in *D*-optimal design theory. Metrika 33, 337–347.

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

Harman, R., Pronzato, L., 2007. Improvements on removing non-optimal support points in *D*-optimum design algorithms. Statistics & Probability Letters 77, 90–94.

Harman, R., Trnovská, M., 2009. Approximate *D*-optimal designs of experiments on the convex hull of a finite set of information matrices. Mathematic Slovaca 59 (5), 693–704.

Khachiyan, L., 1996. Rounding of polytopes in the real number model of computation. Mathematics of Operations Research 21 (2), 307–320.

Khachiyan, L., Todd, M., 1993. On the complexity of approximating the maximal inscribed ellipsoid for a polytope. Math. Programming A61 (2), 137–159.

Kiefer, J., 1974. General equivalence theory for optimum designs (approximate theory). Annals of Stat. 2 (5), 849–879.

Lindsay, B., 1983. The geometry of mixture likelihoods: a general theory. Annals of Statistics 11 (1), 86–94.

Magnus, J., Neudecker, H., 1999. Matrix Differential Calculus, with Applications in Statistics and Econometrics. Wiley, New York.

Mallet, A., 1986. A maximum likelihood estimation method for random coefficient regression models. Biometrika 73 (3), 645–656.

Pázman, A., 1986. Foundations of Optimum Experimental Design. Reidel (Kluwer group), Dordrecht (co-pub. VEDA, Bratislava).

Pukelsheim, F., 1993. Optimal Experimental Design. Wiley, New York.

Schwabe, R., 1996. Optimum Designs for Multi-Factor Models. Springer, New York.

Sibson, R., 1972. Discussion on a paper by H.P. Wynn. Journal of Royal Statistical Society B34, 181–183.

Titterington, D., 1976. Algorithms for computing D-optimal designs on a finite design space. In: Proc. of the 1976 Conference on Information Science and Systems. Dept. of Electronic Engineering, John Hopkins University, Baltimore, pp. 213–216.

Todd, M., Yildirim, E., 2007. On Khachiyan's algorithm for the computation of minimum volume enclosing ellipsoids. Discrete Applied Math. 155, 1731–1744.

Torsney, B., 1983. A moment inequality and monotonicity of an algorithm. In: Kortanek, K., Fiacco, A. (Eds.), Proc. Int. Symp. on Semi-infinite Programming and Applications. Springer, Heidelberg, pp. 249–260.

Wu, C., 1978. Some algorithmic aspects of the theory of optimal designs. Annals of Statistics 6 (6), 1286–1301.

Yu, Y., 2010. Strict monotonicity and convergence rate of Titterington's algorithm for computing D-optimal designs. Comput. Statist. Data Anal. 54, 1419–1425.

Yu, Y., 2011. D-optimal designs via a cocktail algorithm. Stat. Comput. 21, 475–481.