Luc Pronzato, Henry P. Wynn and Anatoly Zhigljavsky,

Extremal measures maximizing functionals based on simplicial volumes

Article (Accepted version) (Refereed)

Original citation:
Pronzato, Luc, Wynn, Henry P. and Zhigljavsky, Anatoly (2016) Extremal measures maximizing functionals based on simplicial volumes. Statistical Papers, 57. pp. 1-17. ISSN 0932-5026
DOI: 10.1007/s00362-016-0767-6

Extremal measures maximizing functionals based on simplicial volumes

Luc Pronzato, Henry Wynn, Anatoly Zhigljavsky

Noname manuscript No.
(will be inserted by the editor)

Extremal measures maximizing functionals based on simplicial volumes

Luc Pronzato · **Henry P. Wynn** · **Anatoly Zhigljavsky**

Received: date / Accepted: date

Abstract We consider functionals measuring the dispersion of a d-dimensional distribution which are based on the volumes of simplices of dimension $k \leq d$ formed by $k+1$ independent copies and raised to some power δ. We study properties of extremal measures that maximize these functionals. In particular, for positive δ we characterize their support and for negative δ we establish connection with potential theory and motivate the application to space-filling design for computer experiments. Several illustrative examples are presented.

Keywords Potential theory · Logarithmic potential · Computer experiments · Space-filling design

Mathematics Subject Classification (2000) 62K05 · 31C15

1 Introduction

Let $\mathscr{X}$ be a compact subset of $\mathbb{R}^d$ and $\mathscr{M}$ be the set of probability measures on the Borel subsets of $\mathscr{X}$. We shall consider the class of functionals $\psi_{k,\delta} : \mathscr{M} \longrightarrow \mathbb{R}^+$ defined by

$$\psi_{k,\delta}(\mu) = \Psi_{k,\delta}(\mu, \ldots, \mu)\,, \tag{1}$$

where

$$\Psi_{k,\delta}(\mu_1, \ldots, \mu_{k+1}) = \int \ldots \int \mathscr{V}_k^{\delta}(x_1, \ldots, x_{k+1})\, \mu_1(\mathrm{d}x_1) \ldots \mu_{k+1}(\mathrm{d}x_{k+1})\,,$$

The work of the first author was partly supported by the ANR project 2011-IS01-001-01 DESIRE (DESIgns for spatial Random fiElds). The third author was supported by the Russian Science Foundation, project Nb. 15-11-30022 ”Global optimization, supercomputing computations, and application”.

Luc Pronzato
CNRS, Laboratoire I3S, UMR 7172, UNS, CNRS; 2000, route des Lucioles, Les Algorithmes, bât. Euclide B, 06900 Sophia Antipolis, France
Tel.: +33-4-92942708
Fax: +33-4-92942896
E-mail: pronzato@cnrs.fr

H.P. Wynn
London School of Economics, Houghton Street, London, WC2A 2AE, UK
E-mail: H.Wynn@lse.ac.uk

A. Zhigljavsky
School of Mathematics, Cardiff University, Senghennydd Road, Cardiff, CF24 4YH, UK
and Lobachevsky Nizhnii Novgorod State University, Russia
E-mail: ZhigljavskyAA@cf.ac.uk

for some δ in $\mathbb{R}$ and $k \in \{1,\dots,d\}$, with $\mathscr{V}_k(x_1,\dots,x_{k+1})$ the volume of the k-dimensional simplex (its length when $k=1$ and area when $k=2$) formed by the $k+1$ vertices $x_1,\dots,x_{k+1}$ in $\mathbb{R}^d$. The volume $\mathscr{V}_k(x_1,\dots,x_{k+1})$ can be computed by the formula

$$\mathscr{V}_k(x_1,\dots,x_{k+1}) = \frac{1}{k!}\,|\det(A)|^{1/2}\,,$$

with

$$A = X^\top X\,,\quad X = [(x_2-x_1)\ (x_3-x_1)\ \cdots (x_{k+1}-x_1)]\,, \tag{2}$$

where the matrix X has size $d\times k$. Define the potential of μ at $x\in\mathbb{R}^d$ by

$$P_{k,\delta,\mu}(x) = \Psi_{k,\delta}(\mu,\dots,\mu,\delta_x)\,, \tag{3}$$

where δ_x is the delta-measure at x and μ appears k times on the right-hand side. Note that $\max_{x\in\mathscr{X}} P_{k,\delta,\mu}(x) \geq \psi_{k,\delta}(\mu)$ for all μ in $\mathscr{M}$ since $\int P_{k,\delta,\mu}(x)\,\mu(\mathrm{d}x) = \psi_{k,\delta}(\mu)$.

The case $\delta=2$ corresponds to an extension of the notion of Wilk's generalized variance and is considered in [11]. In this paper we investigate properties of the functional (1) for general δ.

2 The case $\delta > 0$

When δ is positive we are interested in the maximization of the functional $\psi_{k,\delta}(\mu)$, $\mu\in\mathscr{M}$, and properties of an extremal measure μ^* where the maximum is attained.

2.1 Functionals based on powered distances: $k=1$

For $k=1$, the functional $\psi_{k,\delta}(\cdot)$ defined by (1) corresponds to

$$\psi_{1,\delta}(\mu) = \mathsf{E}\{\|x_1-x_2\|^\delta\}$$

where x_1 and x_2 are supposed to be i.i.d. with the measure μ. Properties of measures that maximize $\psi_{1,\delta}(\mu)$ for $\delta>0$ are investigated in [2]. In particular, it is shown there that for any $\delta>0$ the mass of an optimal measure is concentrated on the boundary of $\mathscr{X}$ and that the support only comprises the extreme points of the convex hull of $\mathscr{X}$ when $\delta>1$. Also, the optimal measure is unique for $\delta<2$; it is supported at no more than $d+1$ points when $\delta>2$.

We can give a more precise statement than in Theorem 2 of [2] for $0<\delta\leq 2$, using the concavity of $\psi_{1,\delta}(\cdot)$, which follows from results discussed in [13] and is based on the fact that $B(\lambda)=\lambda^\alpha$ is a Bernstein function for all $0<\alpha\leq 1$. Indeed, using concavity of $\psi_{1,\delta}(\cdot)$, the measure μ^* is extremal (i.e., it maximizes $\psi_{1,\delta}(\mu)$ with respect to $\mu\in\mathscr{M}$) if and only if the directional derivative

$$F_{\psi_{1,\delta}}(\mu;\nu) = \lim_{\alpha\to 0^+} \frac{\psi_{1,\delta}[(1-\alpha)\mu+\alpha\nu] - \psi_{1,\delta}(\mu)}{\alpha}$$

satisfies $F_{\psi_{1,\delta}}(\mu_k^*;\nu)\leq 0$ for all $\nu\in\mathscr{M}$. Direct calculation gives

$$F_{\psi_{1,\delta}}(\mu;\nu) = 2\left[\int P_{1,\delta,\mu}(x)\,\nu(\mathrm{d}x) - \psi_{1,\delta}(\mu)\right] \tag{4}$$

and we thus obtain the following.

Theorem 1 *For any $0 < \delta \leq 2$, the measure μ^* maximizes $\psi_{1,\delta}(\mu)$ with respect to $\mu \in \mathscr{M}$ if and only if*

$$\max_{x\in\mathscr{X}} P_{1,\delta,\mu^*}(x) = \psi_{1,\delta}(\mu^*)\,.$$

Equivalently, μ^* minimizes $\max_{x\in\mathscr{X}}\,[P_{1,\delta,\mu}(x) - \psi_{1,\delta}(\mu)]$ with respect to $\mu \in \mathscr{M}$.

In connection with the statement of the theorem, we may notice that the extremal measure μ^* does not necessarily minimize $\max_{x\in\mathscr{X}} P_{1,\delta,\mu}(x)$, see [2, Th. 14]. In the next section we show how some of the properties that hold for $k = 1$ can be generalized to the functionals $\psi_{k,\delta}(\cdot)$ with $k \geq 2$.

2.2 Functionals based on powered volumes: $k \geq 2$

2.2.1 A necessary condition for optimality

First note that the existence of an extremal measure follows from the continuity of $\mathscr{V}_k(x_1,\ldots,x_{k+1})$ in each x_i, see [2, Th. 1].

Similarly to the case $k = 1$, we can compute the second order derivative of the functional $\psi_{k,\delta}(\cdot)$. Indeed, for any μ_0, μ_1 in $\mathscr{M}$, we have

$$\begin{aligned}\left.\frac{\partial^2\psi_{k,\delta}[(1-\alpha)\mu_0+\alpha\mu_1]}{\partial\alpha^2}\right|_{\alpha=0} &= k(k+1)\,[\varPsi_{k,\delta}(\mu_0,\ldots,\mu_0,\mu_1,\mu_1) \\ &\quad + \varPsi_{k,\delta}(\mu_0,\ldots,\mu_0) - 2\varPsi_{k,\delta}(\mu_0,\ldots,\mu_0,\mu_1)]\;, \\ &= k(k+1)\int\int P_{k,\delta}(x,y)\,[\mu_0-\mu_1](\mathrm{d}x)[\mu_0-\mu_1](\mathrm{d}y)\,,\end{aligned}$$

where $P_{k,\delta}(x,y) = \int\ldots\int \mathscr{V}_k^\delta(x_1,\ldots,x_{k-1},x,y)\,\mu_0(\mathrm{d}x_1)\ldots\mu_0(\mathrm{d}x_{k-1})$. The proof is by direct calculation, using the symmetry of the kernel $\mathscr{V}_k^\delta(x_1,\ldots,x_{k+1})$ in (1).

For $k = 1$, $P_{1,\delta}(x,y) = \|x-y\|^\delta$, and $\psi_{1,\delta}(\cdot)$ for $\delta \leq 2$ is concave as discussed above. For $\delta = 2$, concavity of $\psi_{k,2}^{1/k}(\cdot)$ is proved in [11] for any $k \in \{1,\ldots,d\}$. We are not aware of any similar result for $k > 1$ and $\delta \neq 2$, so that we have no guarantee that $\psi_{k,\delta}(\cdot)$, even raised to some power less than 1, is concave for $\delta \neq 2$. Therefore, we can only give a necessary condition of optimality for a measure μ^* maximizing $\psi_{k,\delta}(\cdot)$. A similar result for $k = 1$ is Theorem 2 in [2].

Theorem 2 *For any $0 < \delta$, if the measure μ^* maximizes $\psi_{k,\delta}(\mu)$ with respect to $\mu \in \mathscr{M}$, then*

$$\max_{x\in\mathscr{X}} P_{k,\delta,\mu^*}(x) = \psi_{k,\delta}(\mu^*)$$

and $P_{k,\delta,\mu^*}(x) = \psi_{k,\delta}(\mu^*)$ on the support of μ^*.

The proof relies on a straightforward extension of (4) to $k \geq 1$:

$$F_{\psi_{k,\delta}}(\mu;\nu) = (k+1)\left[\int P_{k,\delta,\mu}(x)\,\nu(\mathrm{d}x) - \psi_{k,\delta}(\mu)\right]\,.$$

2.2.2 Support of extremal measures

Below we indicate some properties concerning the support of extremal measures that generalize those in Section 2.1.

Theorem 3 *For any $\delta > \max\{0, k+1-d\}$, the support of any measure μ_k^* maximizing $\psi_{k,\delta}(\mu)$ is a subset of the boundary of $\mathscr{X}$.*

Proof For $\delta > 1$, we can simply use the convexity property of the L_2 norm and multilinearity of the determinant. Indeed, from Binet-Cauchy formula, the squared volume $\mathscr{V}_k^2(x_1,\ldots,x_{k+1})$ can be written as

$$\mathscr{V}_k^2(x_1,\ldots,x_{k+1}) = \frac{1}{(k!)^2} \sum_{1\le i_1<i_2<\cdots<i_k\le d} \det{}^2 \begin{bmatrix} \{x_1\}_{i_1} & \cdots & \{x_{k+1}\}_{i_1} \\ \vdots & \vdots & \vdots \\ \{x_1\}_{i_k} & \cdots & \{x_{k+1}\}_{i_k} \\ 1 & \cdots & 1 \end{bmatrix} . \tag{5}$$

Each determinant in the right-hand side of (5) is linear in x_1, so that, when $\delta > 1$, $\mathscr{V}_k^\delta(x_1,\ldots,x_{k+1})$ is a strictly convex function of x_1. This implies that the potential $P_{k,\delta,\mu_k^*}(x_1)$ is strictly convex in x_1. We then follow similar arguments to those in the proof of [2, Th. 3]. Suppose that x_1 is an interior point of $\mathscr{X}$, and consider a sphere $\mathcal{S}(x_1,r)$ centered at x_1 with radius r included in $\mathscr{X}$. Strict convexity of $P_{k,\delta,\mu_k^*}(\cdot)$ implies that $P_{k,\delta,\mu_k^*}(x_1)$ is strictly smaller than the mean value of $P_{k,\delta,\mu_k^*}(x)$ on $\mathcal{S}(x_1,r)$. From Theorem 2, this mean value is less than or equal to $\psi_{k,\delta}(\mu^*)$, and x_1 cannot be support point of μ_k^*.

For $\delta \le 1$, the proof uses subharmonicity of $P_{k,\delta,\mu_k^*}(\cdot)$ as in [2, Th. 3]. We only need to prove that for fixed $x_2,\ldots,x_{k+1}$, $\mathscr{V}_k^\delta(x_1,\ldots,x_{k+1})$ is a strictly subharmonic function of x_1. From Lemma 2, see Appendix, we have

$$\sum_{i=1}^d \frac{\partial^2 \mathscr{V}_k^\delta(x_1,\ldots,x_{k+1})}{\partial\{x_1\}_i^2} = \delta(\delta+d-k-1)\, \mathscr{V}_k^\delta(x_1,x_2,\ldots,x_{k+1})\, (\mathbf{1}_k^\top A^{-1}\mathbf{1}_k)\,,$$

with A defined in (2) and $\mathbf{1}_k = (1,\ldots,1)^\top \in \mathbb{R}^k$. The right-hand side is strictly positive when $\delta > k+1-d$. $\square$

Theorem 4 *For any $\delta > 1$ and any $k \in \{1,\ldots,d\}$, any measure μ_k^* maximizing $\psi_{k,\delta}(\mu)$ is supported on extreme points of the convex hull of $\mathscr{X}$.*

Proof As shown in the proof of Theorem 3, the potential $P_{k,\delta,\mu_k^*}(x)$ is a strictly convex function of x when $\delta > 1$. Suppose that $x_0 \in \mathscr{X}$ is not an extreme point of the convex hull of $\mathscr{X}$. Then, x_0 can be written as a linear combination of such points z_j with strictly positive weights summing to one. The potential $P_{k,\delta,\mu_k^*}(x_0)$ is then strictly less than the weighted sum of potentials at the z_j, which, from Theorem 2, are all less than or equal to $\psi_{k,\delta}(\mu^*)$. By the same theorem, x_0 cannot be in the support of μ_k^*. $\square$

3 The case $\delta \le 0$

When $\delta < 0$, we are interested in the minimization of the functional $\psi_{k,\delta}(\mu) = \mathsf{E}\{\mathscr{V}_k^\delta(x_1,\ldots,x_{k+1})\}$, $\mu \in \mathscr{M}$. Equivalently, we can consider the maximization of $\psi_{k,\delta}^{1/\delta}(\mu)$, the continuous extension of which at $\delta = 0$ is $\exp\left(\mathsf{E}\{\log[\mathscr{V}_k(x_1,\ldots,x_{k+1})]\}\right)$. We thus define

$$\mathscr{D}_{k,\delta}(\mu) = \begin{cases} \left(\mathsf{E}\{\mathscr{V}_k^\delta(x_1,\ldots,x_{k+1})\}\right)^{1/\delta} & \text{for } \delta \neq 0\,, \\ \exp\left(\mathsf{E}\{\log[\mathscr{V}_k(x_1,\ldots,x_{k+1})]\}\right) & \text{for } \delta = 0\,. \end{cases}$$

The results in Sections 2 have shown that when $\delta > 0$ the support of a measure that maximizes $\mathscr{D}_{k,\delta}$ is sometimes finite and is always included in the boundary of $\mathscr{X}$ when $k \le d-1$. The situation is quite different for $\delta \le 0$, the case we investigate in this section.

In the case $k = 1$, the investigation of the properties of extremal measures $\mu_{1,\delta}^*$ and optimal values $\mathscr{D}_{1,\delta}^* = \mathscr{D}_{1,\delta}(\mu_{1,\delta}^*)$ is one of the main concerns of potential theory, see e.g., [12]. This is equivalent to studying the asymptotic behavior of the so-called Fekete points, defined as follows. Given a natural number n and a real $\delta \leq 0$, the n points $X_n = (x_1, \ldots, x_n) \in \mathscr{X}^n$ are called Fekete points when they maximize

$$\widehat{\mathscr{D}}_{1,\delta}(X_n) = \left[\frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \|x_i - x_j\|^\delta \right]^{1/\delta} \tag{6}$$

for $\delta < 0$ and

$$\widehat{\mathscr{D}}_{1,0}(X_n) = \exp \left\{ \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \log\left(\|x_i - x_j\|\right) \right\} \tag{7}$$

for $\delta = 0$, or equivalently minimize the s-energy, $s = -\delta$, defined by $\mathscr{E}^{(s)}(X_n) = \sum_{1 \leq i < j \leq n} \|x_i - x_j\|^{-s}$ for $s > 0$ and by $\mathscr{E}^{(0)}(X_n) = \sum_{1 \leq i < j \leq n} \log \|x_i - x_j\|^{-1}$ for $s = 0$.

We shall denote by $F_n^{(s)}$ a set of n Fekete points, $s \geq 0$. For instance, when $\mathscr{X} = [-1, 1]$, then the set $F_n^{(0)}$ is uniquely defined and coincides with the zeros of $(1 - x^2)P'_{n-1}(x)$, where P_{n-1} is the Legendre polynomial of degree $n - 1$. One may note that $F_n^{(0)}$ corresponds to the support of a D-optimal design measure for polynomial regression of degree $n - 1$ on $[-1, 1]$, see, e.g., [3, p. 89].

The (logarithmic) transfinite diameter of $\mathscr{X}$ is defined by

$$\tau^{(0)}(\mathscr{X}) = \lim_{n \to \infty} \exp \left\{ -\frac{2}{n(n-1)} \mathscr{E}^{(0)}(F_n^{(0)}) \right\} \tag{8}$$

where the convergence to the limit in (8) is monotonic (in the sense that the exponential term in non-increasing with n). The logarithmic potential associated with $\mu \in \mathscr{M}$ is $P_\mu^{(0)}(z) = \int \log(1/\|z - t\|)\, \mu(\mathrm{d}t)$, the corresponding energy is defined by

$$I^{(0)}(\mu) = \int P_\mu^{(0)}(z)\, \mu(\mathrm{d}z) = \int \int \log \frac{1}{\|z - t\|}\, \mu(\mathrm{d}t)\, \mu(\mathrm{d}z)\,.$$

Similarly, the transfinite diameter of order $s > 0$ is

$$\tau^{(s)}(\mathscr{X}) = \lim_{n \to \infty} \left\{ \frac{2}{n(n-1)} \mathscr{E}^{(s)}(F_n^{(s)}) \right\}^{-1}\,,$$

the s-potential for μ is $P_\mu^{(s)}(z) = \int \|z - t\|^{-s}\, \mu(\mathrm{d}t)$, with associated energy

$$I^{(s)}(\mu) = \int P_\mu^{(s)}(z)\, \mu(\mathrm{d}z) = \int \int \frac{1}{\|z - t\|^s}\, \mu(\mathrm{d}t)\, \mu(\mathrm{d}z)\,.$$

The minimum energy problem involves the determination of

$$I_*^{(s)}(\mathscr{X}) = \inf\{I^{(s)}(\mu) : \mu \in \mathscr{M}\}\,.$$

The logarithmic capacity of $\mathscr{X}$, denoted by $\mathrm{cap}^{(0)}(\mathscr{X})$, is defined by $\mathrm{cap}^{(0)}(\mathscr{X}) = \exp\{-I_*^{(0)}(\mathscr{X})\}$; its s-capacity for $s > 0$ is $\mathrm{cap}^{(s)}(\mathscr{X}) = [I_*^{(s)}(\mathscr{X})]^{-1}$. If $\mathrm{cap}^{(0)}(\mathscr{X}) > 0$, then the extremal measure $\mu_{1,0}^*$ exists with $\mathrm{cap}^{(0)}(\mathscr{X}) = \mathscr{D}_{1,0}(\mu_{1,0}^*)$. Also, for any $s > 0$, if $\mathrm{cap}^{(s)}(\mathscr{X}) > 0$ then $\mu_{1,-s}^*$ exists and $\mathrm{cap}^{(s)}(\mathscr{X}) = [\mathscr{D}_{1,-s}(\mu_{1,-s}^*)]^s$. One of the main results in potential theory is that the capacity of $\mathscr{X}$ coincides with its transfinite diameter: $\mathrm{cap}^{(s)}(\mathscr{X}) = \tau^{(s)}(\mathscr{X})$ for all compact sets $\mathscr{X}$. It also coincides with $\sup_{\mu \in \mathscr{M}} \mathscr{D}_{1,0}(\mu)$ when $s = 0$ and with $\sup_{\mu \in \mathscr{M}} [\mathscr{D}_{1,-s}(\mu)]^s$ when $s > 0$. When $\mathrm{cap}^{(s)}(\mathscr{X}) > 0$, which happens in particular when $\mathscr{X}$ is a compact subset of $\mathbb{R}^d$ and $0 \leq s < d$, then $\mu_{1,-s}^*$ exists, it is called s-energy equilibrium measure and is the weak limit of a sequence of empirical measures associated with Fekete points. Even if $\mathrm{cap}^{(s)}(\mathscr{X}) = 0$ and no measure μ exists with $I^{(s)}(\mu) < \infty$, it is still interesting to study the limiting behaviour of empirical measures of Fekete points, see [4].

Fekete point are extremely difficult to construct, except for a few particular cases. When $s = 0$, Fekete points necessarily lie on $\partial_\infty(\mathscr{X})$, the outer boundary of $\mathscr{X}$. This implies that the extreme (equilibrium) measure $\mu^*_{1,0}$ is supported on $\partial_\infty(\mathscr{X})$ too. Consequently, $\mathrm{cap}^{(0)}(\mathscr{X}) = \mathrm{cap}^{(0)}(\partial_\infty(\mathscr{X}))$. If the outer boundary $\partial_\infty(\mathscr{X})$ is a continuum, then $\mathrm{supp}(\mu^*_{1,0}) = \partial_\infty(\mathscr{X})$. In general, $\partial_\infty(\mathscr{X}) \setminus \mathrm{supp}(\mu^*_{1,0})$ has capacity zero.

Example 1: $d = 1$, $\mathscr{X} = [0, 1]$. The extremal measure $\mu^*_{1,0}$ has the arcsine density

$$\pi_0(t) = \frac{1}{\pi\sqrt{t(1-t)}}$$

on $[0, 1]$ and $\mathrm{cap}^{(0)}(\mathscr{X}) = 1/4$. More generally, the measure $\mu^*_{1,\delta}$ maximizing $\mathscr{D}_{1,\delta}(\mu)$ with $\delta \in (-1, 0]$ corresponds to the Beta distribution on $[0, 1]$ with density

$$\pi_\delta(t) = \frac{1}{B[(1-\delta)/2, (1-\delta)/2]} \, \frac{1}{\sqrt{[t(1-t)]^{\delta+1}}} \,,$$

see, e.g., [14]. This distribution is uniform for $\delta = -1$, with $\mathscr{E}^{(0)}(F_n^{(0)})$ growing as $n^2 \log n$, and, as mentioned in [4], the limiting distribution of Fekete points is uniform for every $\delta \le -1$.

Example 2: $\mathscr{X} = \mathscr{B}_d(\mathbf{0}, \rho)$. As indicated in [4], the extremal measure $\mu^*_{1,\delta}$ maximizing $\mathscr{D}_{1,\delta}(\cdot)$ is uniquely defined for $-d < \delta \le 0$ (as the $|\delta|$-energy equilibrium measure). From [6, p. 163], $-d < \delta < 2 - d$, it has the density

$$\varphi_\delta(x) = \frac{C}{(\rho^2 - \|x\|^2)^{(d+\delta)/2}} \,,\ x \in \mathscr{B}_d(\mathbf{0}, \rho) \,,$$

where $C = R^\delta \pi^{-d/2} \Gamma(1 - \delta/2)/\Gamma(1 - (d + \delta)/2)$. For $2 - d \le \delta \le 0$, $\mu^*_{1,\delta}$ is uniform on the sphere $\mathcal{S}_d(\mathbf{0}, \rho)$. For $\delta \le -d$, any sequence of Fekete points is asymptotically uniformly distributed in $\mathscr{B}_d(\mathbf{0}, \rho)$, with $\mathscr{E}^{(-\delta)}(F_n^{(-\delta)})$ growing as $n^2 \log n$ for $\delta = -d$ and as $n^{1-\delta/d}$ for $\delta < -d$, see [4].

To the best of our knowledge, no theory is available which would cover the case $k > 1$. In the next section we only present results concerning a particular example which illustrate the difference with the case $k = 1$.

4 Particular case: $\mathscr{X} = \mathscr{B}_d(\mathbf{0}, \rho)$

Take $\mathscr{X} = \mathscr{B}_d(\mathbf{0}, \rho)$, the closed ball of $\mathbb{R}^d$ centered at the origin $\mathbf{0}$ with radius ρ.

Case $\delta = 2$. Let μ_0 be the uniform measure on the sphere $\mathcal{S}_d(\mathbf{0}, \rho)$ (the boundary of $\mathscr{B}_d(\mathbf{0}, \rho)$). Then, the covariance matrix $V_{\mu_0} = \int xx^\top \mu_0(\mathrm{d}x)$ is proportional to the identity matrix I_d, $V_{\mu_0} = \rho^2 I_d/d$. Take $k = d$. We have

$$\max_{x \in \mathscr{X}} x^\top \nabla_{\psi_{d,2}}[V_{\mu_0}] x = \frac{(d+1)\rho^{2d}}{d^{d-1} d!} = \mathrm{trace}\{V_{\mu_0} \nabla_{\psi_{d,2}}[V_{\mu_0}]\} \,,$$

where $\nabla_{\psi_{d,2}}[V_\mu] = [(d+1)/d!] \det(V_\mu) V_\mu^{-1}$ is the gradient of $\psi_{d,2}(\mu)$ considered as a function of V_μ, see [11]. From Theorem 4.1 in the same paper, this implies that μ_0 maximizes $\psi_{d,2}(\mu)$.

Let μ_d be the measure that allocates mass $1/(d+1)$ at each vertex of a d regular simplex having its $d+1$ vertices on $\mathcal{S}_d(\mathbf{0},\rho)$, with squared volume $\rho^{2d}(d+1)^{d+1}/[d^d(d!)^2]$. We also have $V_{\mu_d}=\rho^2 I_d/d$, so that μ_d also maximizes $\psi_{d,2}(\cdot)$. In view of [11, Remark 4.2], μ_0 and μ_d maximize $\psi_{k,2}$ for all k in $\{1,\ldots,d\}$.

Let now μ_k be the measure that allocates mass $1/(k+1)$ at each vertex of a k regular simplex $\mathscr{P}_k$, centered at the origin, with its vertices on $\mathcal{S}_d(\mathbf{0},\rho)$. The squared volume of $\mathscr{P}_k$ equals $\rho^{2k}\,(k+1)^{k+1}/[k^k(k!)^2]$. Without any loss of generality, we can choose the orientation of the space so that V_{μ_k} is diagonal, with its first k diagonal elements equal to ρ^2/k and the other elements equal to zero. Note that $\psi_{k',2}(\mu_k)=0$ for $k'>k$. Direct calculations give

$$\psi_{k,2}(\mu_k)=\frac{k+1}{k!}\,\frac{\rho^{2k}}{k^k}\leq\psi_k(\mu_0)=\frac{k+1}{k!}\binom{d}{k}\frac{\rho^{2k}}{d^k}\,,$$

with equality for $k=1$ and $k=d$, the inequality being strict otherwise. Figure 1 presents the efficiency $[\psi_{k,2}(\mu_k)/\psi_{k,2}(\mu_0)]^{1/k}$ as a function of k when $d=20$.

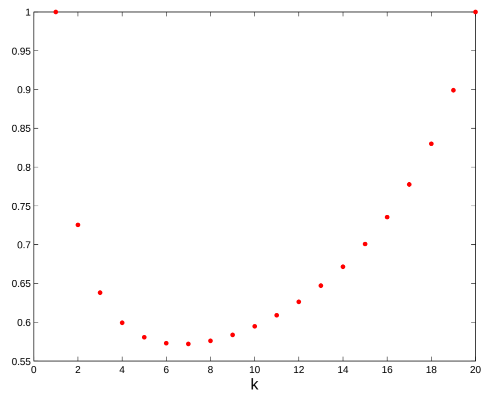

Fig. 1 Efficiency $[\psi_{k,2}(\mu_k)/\psi_{k,2}(\mu_0)]^{1/k}$ as a function of k when $d=20$

Case $\delta>2$. We can show that for any $\delta>2$ the measure μ maximizes $\psi_{d,\delta}(\cdot)$ if and only if it coincides with one of the measures μ_d introduces above.

The proof follows closely that of Theorem 7 in [2] which concerns the case $k=1$. We have

$$\begin{aligned}\psi_{d,\delta}(\mu)&=\int \mathscr{V}_d^{\delta-2}(x_1,\ldots,x_{d+1})\,\mathscr{V}_d^2(x_1,\ldots,x_{d+1})\,\mu(\mathrm{d}x_1)\ldots\mu(\mathrm{d}x_{d+1})\\&\leq\max_{x_1,\ldots,x_{d+1}}\mathscr{V}_d^{\delta-2}(x_1,\ldots,x_{d+1})\int\mathscr{V}_d^2(x_1,\ldots,x_{d+1})\,\mu(\mathrm{d}x_1)\ldots\mu(\mathrm{d}x_{d+1})\,.\end{aligned}\qquad(9)$$

Since $\mathscr{V}_d^*=\max_{x_1,\ldots,x_{d+1}}\mathscr{V}_d(x_1,\ldots,x_{d+1})=\rho^d(d+1)^{(d+1)/2}/[d^{d/2}\,d!]$ and the uniform measure μ_0 on the sphere $\mathcal{S}_d(\mathbf{0},\rho)$ is extremal for $\psi_{d,2}(\cdot)$, we get

$$\psi_{d,\delta}(\mu)\leq\left(\frac{\rho^{2d}(d+1)^{d+1}}{d^d(d!)^2}\right)^{\delta/2-1}\psi_{d,2}(\mu_0)=\rho^{d\delta}\,\frac{(d+1)^{(d+1)\delta/2-d}}{(d!)^{\delta-1}\,d^{d\delta/2}}\,.$$

On the other hand, this is exactly the value $\psi_{d,\delta}(\mu_d)$. Therefore, for the measure μ to be extremal we need to have equality in (9), which requires that $\mathscr{V}_d(x_1,\ldots,x_{d+1})=\mathscr{V}_d^*$ for all $(k+1)$-tuples that contribute to the integral. This forces the extremal measure to have the form indicated.

Consider the case $d = 2$, $\rho = 1$. Figure 2 presents the potential $P_{2,\delta,\mu_2}(x(t))$ with $x(t) = (\cos(t), \sin(t))$ as a function of $t \in [0, 2\pi]$ for $\delta = 1$ (left) and $\delta = 4$ (right), with μ_2 allocating weight $1/3$ at each of the three points $(1, 0)$, $(\cos(2\pi/3), \sin(2\pi/3))$ and $(\cos(4\pi/3), \sin(4\pi/3))$. The value of $\psi_{2,\delta}(\mu_2)$ is indicated in dashed line. The figure illustrates the fact that μ_2 is extremal for $\psi_{2,4}(\cdot)$ but is not extremal for $\psi_{2,1}(\cdot)$ since the necessary condition of Theorem 2 is violated. The analytic forms for the potentials are $P_{2,1,\mu_2}(x(t)) = (\sqrt{3}/18)+(\sqrt{3}/9)\cos(t)+(1/3)\sin(t)$ for $0 \le t \le 2\pi/3$ and $P_{2,4,\mu_2}(x(t)) = 57/128 + (3/16)\cos(3t)$ for $0 \le t \le 2\pi$.

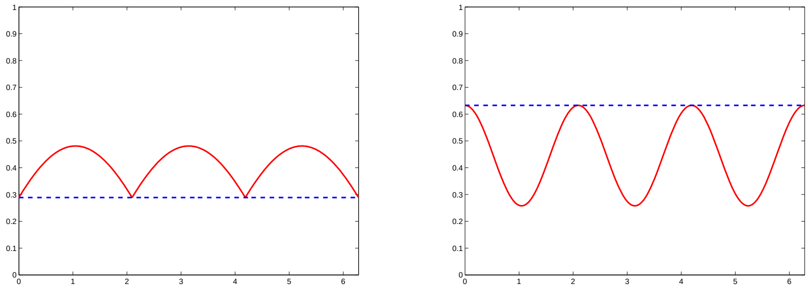

Fig. 2 Potential $P_{2,\delta,\mu_2}(x(t))$, with $x(t) = (\cos(t), \sin(t))$, as a function of $t \in [0, 2\pi]$ (solid line) and value of $\psi_{2,\delta}(\mu_2)$ (dashed line) for $\delta = 1$ (left) and $\delta = 4$ (right); μ_2 allocates weight $1/3$ at each point of an equilateral triangle with vertices on $\mathscr{S}_2(\mathbf{0}, 1)$

Uniform measure on the circle $\mathscr{S}_2(\mathbf{0}, 1)$. Assume that $k = d = 2$, $\mathscr{X} = \mathscr{B}(\mathbf{0}, 1)$, and consider the uniform measure $\mu_{\mathscr{S}}$ on $\mathscr{S}_2(\mathbf{0}, 1)$, which is optimal for $\delta = 2$.

Consider n-point sets X_n containing the points $x_j = (\cos(2\pi j/n), \sin(2\pi j/n))$, $j = 0, \ldots, n-1$, with empirical measure converging to $\mu_{\mathscr{S}}$. The empirical version of (1) is

$$\psi_{2,\delta}(X_n) = \frac{2}{(n-1)(n-2)} \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \mathscr{V}_2^\delta(x_0, x_i, x_j)\,.$$

Direct calculations give

$$\begin{aligned}
\psi_{2,1}(X_n) &= \frac{3\,n}{2(n-1)(n-2)} \cot(\pi/n) = \frac{3}{2\,\pi}\left(1 + \frac{3}{n} + O\left(n^{-2}\right)\right) \\
\psi_{2,2}(X_n) &= \frac{3\,n^2}{2^3(n-1)(n-2)} \\
\psi_{2,3}(X_n) &= \frac{35}{32\,\pi}\left(1 + \frac{3}{n} + O\left(n^{-2}\right)\right) \\
\psi_{2,4}(X_n) &= \frac{45\,n^2}{2^7\,(n-1)(n-2)} \\
\psi_{2,5}(X_n) &= \frac{3003}{2560\,\pi}\left(1 + \frac{3}{n} + O\left(n^{-2}\right)\right) \\
\psi_{2,6}(X_n) &= \frac{105\,n^2}{2^8\,(n-1)(n-2)} \\
\psi_{2,8}(X_n) &= \frac{17325\,n^2}{2^{15}\,(n-1)(n-2)}
\end{aligned}$$

Figure 3 presents $\psi_{2,\delta}(\mu_{\mathscr{S}})$ as a function of $\delta \in [0, 8]$. The stars indicate the exact values obtained from the expressions above.

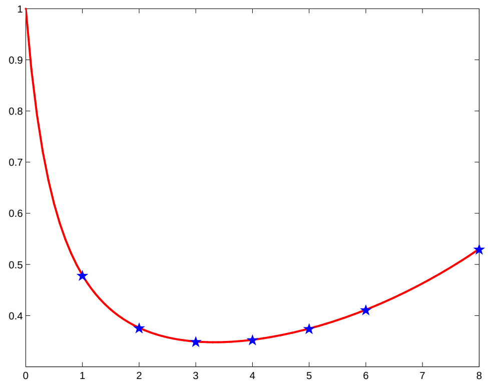

Fig. 3 $\psi_{2,\delta}(\mu_{\mathcal{S}})$ as a function of $\delta \in [0,8]$, for $\mu_{\mathcal{S}}$ uniform on $\mathcal{S}_2(\mathbf{0},1)$

By considering the potential $P_{2,\delta,\mu_{\mathcal{S}}}(\cdot)$ at the origin $\mathbf{0}$ for δ close to zero, we can show that the necessary condition of Theorem 2 for $\mu_{\mathcal{S}}$ being optimal is violated for $\delta < 0$. Indeed, we have

$$\psi_{2,\delta}(\mu_{\mathcal{S}}) = 1-(2\log 2)\,\delta+c_1\delta^2+O(\delta^3)\,, \quad P_{2,\delta,\mu_{\mathcal{S}}}(\mathbf{0}) = 1-(2\log 2)\,\delta+c_2\delta^2+O(\delta^3)\,,$$

with $c_1 \simeq 2.1946$ and $c_2 \simeq 1.3721$, so that $P_{2,\delta,\mu_{\mathcal{S}}}(\mathbf{0}) < \psi_{2,\delta}(\mu_{\mathcal{S}})$ for all $\delta \neq 0$. However, for negative δ, $\psi_{2,\delta}(\cdot)$ should be minimized, the necessary condition for optimality of μ^* becomes $P_{2,\delta,\mu^*}(x) \geq \psi_{2,\delta}(\mu^*)$ for any $x \in \mathscr{B}(\mathbf{0},1)$, and is thus violated for $\mu_{\mathcal{S}}$ at $x = \mathbf{0}$. Although $\mu_{\mathcal{S}}$ is not optimal for negative δ, $\psi_{2,\delta}(\mu_{\mathcal{S}})$ remains finite for $\delta > -2/3$. If $\mathscr{X}$ is reduced to the circle $\mathcal{S}_2(\mathbf{0},1)$, then the n-point sets X_n are Fekete points (in the usual sense, for $k=1$) and can be considered as generalized Fekete points for $k=2$. One can show that $\psi_{2,\delta}(X_n) = O(n^{-(2+3\delta)})$ for $\delta < -2/3$.

On the other hand, for $k=1$, the measure $\mu_{\mathcal{S}}$ is optimal for $0 \leq \delta \leq 2$ and $\psi_{1,\delta}(\mu_{\mathcal{S}})$ is finite for all $\delta > -1$; $\lim_{n\to\infty} \mathscr{E}^{(1)}(X_n)/(n^2 \log n) = 1$ and $\mathscr{E}^{(-\delta)}(X_n)$ grows like $n^{1-\delta}$ ($\psi_{1,\delta}(X_n)$ grows like $n^{-(1+\delta)}$) for $\delta < -1$.

5 Generalized Fekete points and design criteria for computer experiments

For a n-point sample, or design, $X_n = \{x_1,\ldots,x_n\}$, $n \geq k+1$, as extensions of (6) and (7), we define

$$\widehat{\mathscr{D}}_{k,\delta}(X_n) = \left[\binom{n}{k+1}^{-1} \sum_{1\leq j_1<j_2<\cdots<j_{k+1}\leq n} \mathscr{V}_k^\delta(x_{j_1},\ldots,x_{j_{k+1}})\right]^{1/\delta}\,,\ \delta \neq 0\,,$$

and

$$\widehat{\mathscr{D}}_{k,0}(X_n) = \exp\left\{\binom{n}{k+1}^{-1} \sum_{1\leq j_1<j_2<\cdots<j_{k+1}\leq n} \log\left[\mathscr{V}_k(x_{j_1},\ldots,x_{j_{k+1}})\right]\right\}\,.$$

The functions $\widehat{\mathscr{D}}_{1,\delta}(\cdot)$ with $\delta \leq 0$ have been suggested as criteria to be maximized for the construction of space-filling designs for computer experiments. An optimal design $X^*_{n,1,\delta}$ maximizing $\widehat{\mathscr{D}}_{1,\delta}(X_n)$ is a set of Fekete points, as defined in Section 3. In particular, $\widehat{\mathscr{D}}_{1,-2}(\cdot)$ corresponds to the energy criterion considered in [1]; see also [8,9].

Lemma 1 *Take $\delta \leq 0$, $k \in \{1, \ldots, d\}$, and consider a design X_n with $\widehat{\mathscr{D}}_{k,\delta}(X_n) > 0$. Then, for any $k' \in \{1, \ldots, k\}$, the projection of X_n on any $(d+1-k')$-dimensional linear subspace contains at least $\lfloor n/k' \rfloor + n (\mathrm{mod}\, k')$ distinct elements.*

Proof Take $k' \in \{1, \ldots, k\}$, any $(k'-1)$-dimensional subspace of $\mathbb{R}^d$ contains k' points at most since otherwise one could find $k+1$ points in the same $(k-1)$-dimensional subspace, contradicting the property $\widehat{\mathscr{D}}_{k,\delta}(X_n) > 0$. Consider the projection p_i of one point x_i of X_n on a $(d+1-k')$-dimensional linear subspace. There are necessarily k' points at most in X_n, including x_i itself, that yield the same projection p_i. $\square$

One may notice the difference with the usual projection properties considered in design for computer experiments, where only projections onto fixed canonical subspaces are considered. For instance, Latin hypercube design [7] ensures that all projections on coordinate axes have exactly n points; however, it does not protect against all points lying on a single line.

Letting δ tend to $-\infty$ in $\widehat{\mathscr{D}}_{1,\delta}(\cdot)$ yields maximin-distance optimal design, see [5], equivalent to the solution of a sphere-packing problem. More generally, for a given sample X_n, we define

$$\widehat{\mathscr{D}}_{k,-\infty}(X_n) = \min_{1 \leq j_1 < j_2 < \cdots < j_{k+1} \leq n} \mathscr{V}_k(x_{j_1}, \ldots, x_{j_{k+1}})\,. \tag{10}$$

Then, $\widehat{\mathscr{D}}_{k,-\infty}(X_n) \leq \widehat{\mathscr{D}}_{k,\delta}(X_n)$ for any $\delta \in \mathbb{R}$, with $\lim_{\delta \to -\infty} \widehat{\mathscr{D}}_{k,\delta}(X_n) = \widehat{\mathscr{D}}_{k,-\infty}(X_n)$. Also, if $X^*_{n,k,\delta} \in \mathscr{X}^n$ maximizes $\widehat{\mathscr{D}}_{k,\delta}(\cdot)$ and $X^*_{n,k,-\infty} \in \mathscr{X}^n$ is a maximin-optimal design that maximizes $\widehat{\mathscr{D}}_{k,-\infty}(\cdot)$, then we have the following bound on the maximin-efficiency of $X^*_{n,k,\delta}$,

$$\frac{\widehat{\mathscr{D}}_{k,-\infty}(X^*_{n,k,\delta})}{\widehat{\mathscr{D}}_{k,-\infty}(X^*_{n,k,-\infty})} \geq \binom{n}{k+1}^{1/\delta},$$

see [10, Chap. 8]. In general, $\widehat{\mathscr{D}}_{1,\delta}(\cdot)$ with δ not too small is easier to optimize than $\widehat{\mathscr{D}}_{1,-\infty}(\cdot)$, see, e.g., [1,8]; one may expect the same to be true for $k > 1$. Notice that from the discussion in Section 3, it is recommended to choose $\delta \leq -d$ to obtain designs evenly spread over $\mathscr{X}$ when maximizing $\widehat{\mathscr{D}}_{1,\delta}(\cdot)$. Also note that, contrary to $\widehat{\mathscr{D}}_{1,-\infty}(X_n)$ which only depends on the relative distances between neighboring pairs of points, the value of $\widehat{\mathscr{D}}_{k,-\infty}(X_n)$ with $k > 1$ is influenced by the respective positions of points whatever their relative distances, see Lemma 1.

Example 3. We report the maximin optimal designs we have calculated for values of n between 5 and 8 for $d = 2$ and $\mathscr{X} = [0,1]^2$. Note that we have in fact equivalence classes of optimal designs, considering symmetries ($\{x\}_i \mapsto 1 - \{x\}_i$, $i = 1, 2$) and a permutation of coordinates; only one representant is indicated. We represent designs as matrices, with column i corresponding to coordinates of the i-th design point. Maximin-distance optimal designs ($k = 1$) can be found for instance at `http://www.packomania.com/`. We have used the following procedure to determine maximin optimal designs for $\widehat{\mathscr{D}}_{2,-\infty}(\cdot)$: ($i$) a global random search algorithm, initialized at a random Latin hypercube design, generates a first design $X_n^{(1)}$; (ii) a local maximization (subgradient-type method, see Appendix) initialized at $X_n^{(1)}$, generates a second design $X_n^{(2)}$; (iii) the configuration of the best design obtained after several repetitions of steps (i) and (ii) is used to determine analytically the optimal design having this configuration. Although we only proved local optimality, we conjecture that the designs presented are indeed optimal for $\widehat{\mathscr{D}}_{2,-\infty}(\cdot)$.

The maximin-distance optimal design ($k=1$) with $n=5$ points is

$$X^*_{5,1,-\infty} = \begin{bmatrix} 0 & 1 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 1 & 1/2 \end{bmatrix},$$

with $\widehat{\mathscr{D}}_{1,-\infty}(X^*_{5,1,-\infty}) = \sqrt{2}/2 \simeq 0.70711$. For $k=2$, we get $\widehat{\mathscr{D}}_{2,-\infty}(X^*_{5,1,-\infty}) = 0$ since the presence of a central point produces two alignements of three points. On the other hand, the optimal design that we have obtained for $\widehat{\mathscr{D}}_{2,-\infty}(\cdot)$ is

$$X^*_{5,2,-\infty} = \begin{bmatrix} 1/3 & 1 & 1 & 1-\sqrt{3}/3 & 0 \\ 0 & 0 & 2/3 & 1 & \sqrt{3}/3 \end{bmatrix},$$

with $\widehat{\mathscr{D}}_{1,-\infty}(X^*_{5,2,-\infty}) = \sqrt{2}(1-\sqrt{3}/3) \simeq 0.59771$ and $\widehat{\mathscr{D}}_{2,-\infty}(X^*_{5,2,-\infty}) = \sqrt{3}/9 \simeq 0.19245$.

For $n=6$, there exists a continuum of maximin optimal designs $X^*_{6,2,-\infty}$, of the form

$$X^*_{6,2,-\infty} = \begin{bmatrix} 1/2 & 1 & 1 & 1/2 & 0 & 0 \\ 0 & 1/2-a & 1-a & 1 & 1/2+a & a \end{bmatrix},\ a \in [0,1/2],$$

all with $\widehat{\mathscr{D}}_{2,-\infty}(X^*_{6,2,-\infty}) = 1/8$. Notice that $X^*_{5,2,-\infty}$ and $X^*_{6,2,-\infty}$ do not contain any central point.

For $n=7$, we have obtained

$$X^*_{7,2,-\infty} = \begin{bmatrix} 0 & 2/3 & 1 & 1 & 2/3 & 0 & 1/6 \\ 0 & 0 & 1/4 & 3/4 & 1 & 1 & 1/2 \end{bmatrix},$$

with $\widehat{\mathscr{D}}_{2,-\infty}(X^*_{7,2,-\infty}) = 1/12 \simeq 0.08333$.

The maximin optimal design for $k=2$ and $n=8$ is

$$X^*_{8,2,-\infty} = \begin{bmatrix} a & 1 & 1 & 1-a & 0 & 0 & c & 1-c \\ 0 & 0 & 1-b & 1 & 1 & b & 1-b & b \end{bmatrix},$$

with $a = (7-\sqrt{13})/18$, $b = (5-\sqrt{13})/6$ and $c = (7-\sqrt{13})/9$, with $\widehat{\mathscr{D}}_{2,-\infty}(X^*_{8,2,-\infty}) = (1+\sqrt{13})(7-\sqrt{13})/216 \simeq 0.072376$.

The designs $X^*_{5,2,-\infty}$ to $X^*_{8,2,-\infty}$ are presented on Figure 4. The circles centered at the design points have radius $r_n = \widehat{\mathscr{D}}_{1,-\infty}(X_n)/2$, with $r_n < \widehat{\mathscr{D}}_{1,-\infty}(X^*_{n,1,-\infty})/2$ since the designs X_n are not maximin-distance optimal. On the other hand, any triplet of design points forms a triangle with area at least $\widehat{\mathscr{D}}_{2,-\infty}(X^*_{n,2,-\infty})$. Note that for each n equality is achieved for several triplets of points. For instance, when $n=5$, the area of the four triangles ABE, ADE, CDE and BCD on Figure 4-top-left equals $\widehat{\mathscr{D}}_{2,-\infty}(X^*_{5,2,-\infty}) = \sqrt{3}/9$, and any other five-point design contains a triangle with area $\mathcal{A} \leq \sqrt{3}/9$.

Appendix

Lemma 2 *Consider matrix A given by (2). The Laplacian of $\det^\alpha(A)$ considered as a function of x_1 is*

$$\sum_{i=1}^{d} \frac{\partial^2 \det^\alpha(A)}{\partial\{x_1\}_i^2} = 2\alpha(2\alpha+d-k-1)\det^\alpha(A)\,(\mathbf{1}_k^\top A^{-1}\mathbf{1}_k)\,, \tag{11}$$

where $\mathbf{1}_k = (1,\ldots,1)^\top \in \mathbb{R}^k$.

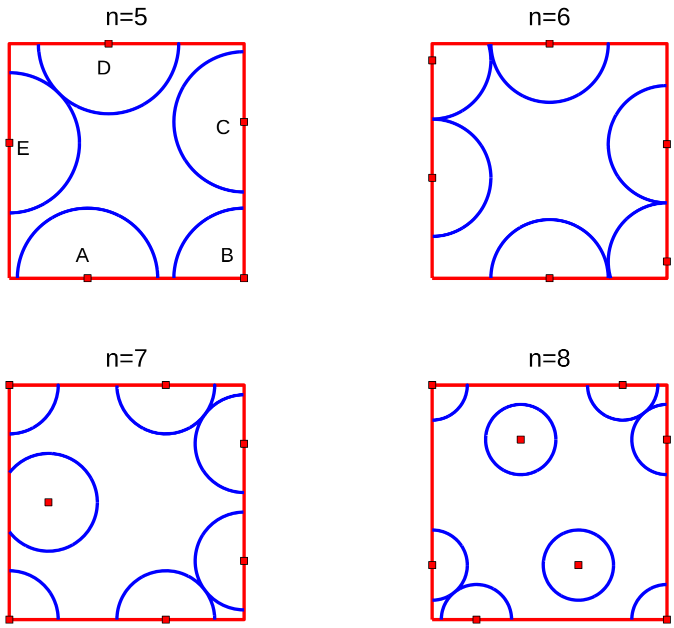

Fig. 4 Optimal designs for $\widehat{\mathscr{D}}_{2,-\infty}(\cdot)$ for n from 5 to 8; $a = 3/7$ in $X^*_{6,2,-\infty}$; the circles have radius $\widehat{\mathscr{D}}_{1,-\infty}(X_n)/2$

Proof We have

$$\begin{aligned}
\frac{\partial \det(A)}{\partial \{x_1\}_i} &= \det(A)\,\mathrm{trace}\left(A^{-1}\frac{\partial A}{\partial \{x_1\}_i}\right)\\
\frac{\partial^2 \det(A)}{\partial \{x_1\}_i^2} &= -\det(A)\,\mathrm{trace}\left(A^{-1}\frac{\partial A}{\partial \{x_1\}_i}A^{-1}\frac{\partial A}{\partial \{x_1\}_i}\right)\\
&\qquad + \det(A)\,\mathrm{trace}^2\left(A^{-1}\frac{\partial A}{\partial \{x_1\}_i}\right) + \det(A)\,\mathrm{trace}\left(A^{-1}\frac{\partial^2 A}{\partial \{x_1\}_i^2}\right),
\end{aligned}$$

where $\partial A/\partial\{x_1\}_i = -[\mathbf{1}_k\Delta_i^\top + \Delta_i\mathbf{1}_k^\top]$ and $\partial^2 A/\partial\{x_1\}_i^2 = 2\mathbf{1}_k\mathbf{1}_k^\top$, with $\Delta_i = (\{x_2 - x_1\}_i, \ldots, \{x_{k+1} - x_1\}_i)^\top \in \mathbb{R}^k$. This gives

$$\frac{\partial^2 \det(A)}{\partial \{x_1\}_i^2} = 2\,\det(A)\,\left\{\mathbf{1}_k^\top A^{-1}\mathbf{1}_k(1 - \Delta_i^\top A^{-1}\Delta_i) + (\mathbf{1}_k^\top A^{-1}\Delta_i)^2\right\}.$$

Noting that $\sum_{i=1}^d \Delta_i\Delta_i^\top = A$, we have $\sum_{i=1}^d \Delta_i^\top A^{-1}\Delta_i = \mathrm{trace}(I_k) = k$ and obtain

$$\begin{aligned}
\sum_{i=1}^d \left(\frac{\partial \det(A)}{\partial \{x_1\}_i}\right)^2 &= \det{}^2(A)\sum_{i=1}^d \mathrm{trace}^2\left(A^{-1}\frac{\partial A}{\partial \{x_1\}_i}\right)\\
&= \det{}^2(A)\sum_{i=1}^d \mathrm{trace}^2\left(A^{-1}[\mathbf{1}_k\Delta_i^\top + \Delta_i\mathbf{1}_k^\top]\right)\\
&= 4\det{}^2(A)\,\mathbf{1}_k^\top A^{-1}\mathbf{1}_k
\end{aligned}$$

and

$$\sum_{i=1}^d \frac{\partial^2 \det(A)}{\partial \{x_1\}_i^2} = 2\det(A)\,\mathbf{1}_k^\top A^{-1}\mathbf{1}_k\,(d + 1 - k)\,.$$

Now,

$$\frac{\partial \det^\alpha(A)}{\partial\{x_1\}_i} = \alpha \det^{\alpha-1}(A)\, \frac{\partial \det(A)}{\partial\{x_1\}_i}$$

$$\frac{\partial^2 \det^\alpha(A)}{\partial\{x_1\}_i^2} = \alpha(\alpha-1)\det^{\alpha-2}(A) \left(\frac{\partial \det(A)}{\partial\{x_1\}_i}\right)^2 + \alpha \det(A)^{\alpha-1}\, \frac{\partial^2 \det(A)}{\partial\{x_1\}_i^2}\,,$$

which finally gives (11). □

A subgradient-type algorithm to maximize $\widehat{\mathscr{D}}_{k,-\infty}(\cdot)$.

Consider a design $X_n = (x_1, \ldots, x_n)$, with each $x_i \in \mathscr{X}$, a convex subset of $\mathbb{R}^d$, as a vector in $\mathbb{R}^{n\times d}$. The function $\widehat{\mathscr{D}}_{k,-\infty}(\cdot)$ defined in (10) is not concave (due to the presence of min), but is Lipschitz and thus differentiable almost everywhere. At points X_n where it fails to be differentiable, we consider any particular gradient from the subdifferential,

$$\nabla \widehat{\mathscr{D}}_{k,-\infty}(X_n) = \nabla v_{j_1,\ldots,j_{k+1}}(X_n)$$

where $x_{j_1}, \ldots, x_{j_{k+1}}$ are such that $\mathscr{V}_k(x_{j_1}, \ldots, x_{j_{k+1}}) = \widehat{\mathscr{D}}_{k,-\infty}(X_n)$ and where $\nabla v_{j_1,\ldots,j_{k+1}}(X_n)$ denotes the usual gradient of the function $\mathscr{V}_k(x_{j_1}, \ldots, x_{j_{k+1}})$. Our subgradient-type algorithm then corresponds to the following sequence of iterations, where the current design $X_n^{(t)}$ is updated into

$$X_n^{(t+1)} = P_{\mathscr{X}}\left[X_n^{(t)} + \gamma_t \nabla \widehat{\mathscr{D}}_{k,-\infty}(X_n^{(t)})\right],$$

where $P_{\mathscr{X}}[\cdot]$ denotes the orthogonal projection on $\mathscr{X}$ and $\gamma_t > 0$, $\gamma_t \searrow 0$, $\sum_t \gamma_t = \infty$, $\sum_t \gamma_t^2 < \infty$.

Direct calculation gives

$$\frac{\partial v_{j_1,\ldots,j_{k+1}}(X_n)}{\partial\{x_j\}_\ell} = \begin{cases} 0 & \text{if } j \notin \{j_1, \ldots, j_{k+1}\} \\ \frac{1}{2k!}\, \det^{1/2}(A_{j_1,\ldots,j_{k+1}})\, \mathrm{trace}\left[A^{-1}_{j_1,\ldots,j_{k+1}} \frac{\partial A_{j_1,\ldots,j_{k+1}}}{\partial\{x_j\}_\ell}\right] & \text{otherwise,} \end{cases}$$

where

$$A_{j_1,\ldots,j_{k+1}} = \left(\begin{bmatrix} (x_{j_2} - x_{j_1})^\top \\ (x_{j_3} - x_{j_1})^\top \\ \vdots \\ (x_{j_{k+1}} - x_{j_1})^\top \end{bmatrix} \begin{bmatrix} (x_{j_2} - x_{j_1}) & (x_{j_3} - x_{j_1}) & \cdots & (x_{j_{k+1}} - x_{j_1}) \end{bmatrix} \right),$$

so that

$$\mathrm{trace}\left[A^{-1}_{j_1,\ldots,j_{k+1}} \frac{\partial A_{j_1,\ldots,j_{k+1}}}{\partial\{x_j\}_\ell}\right] = 2 \left\{ A^{-1}_{j_1,\ldots,j_{k+1}} \begin{bmatrix} \{(x_{j_2} - x_{j_1})\}_\ell \\ \{(x_{j_3} - x_{j_1})\}_\ell \\ \vdots \\ \{(x_{j_{k+1}} - x_{j_1})\}_\ell \end{bmatrix} \right\}_{j-1}$$

for $j \in \{j_1, \ldots, j_{k+1}\}$, $j \neq j_1$, and

$$\mathrm{trace}\left[A^{-1}_{j_1,\ldots,j_{k+1}} \frac{\partial A_{j_1,\ldots,j_{k+1}}}{\partial\{x_{j_1}\}_\ell}\right] = -2 \sum_{i=1}^{k} \left\{ A^{-1}_{j_1,\ldots,j_{k+1}} \begin{bmatrix} \{(x_{j_2} - x_{j_1})\}_\ell \\ \{(x_{j_3} - x_{j_1})\}_\ell \\ \vdots \\ \{(x_{j_{k+1}} - x_{j_1})\}_\ell \end{bmatrix} \right\}_i.$$

References

1. P. Audze and V. Eglais. New approach for planning out of experiments. *Problems of Dynamics and Strengths*, 35:104–107, 1977.
2. G. Björck. Distributions of positive mass, which maximize a certain generalized energy integral. *Arkiv för Matematik*, 3(21):255–269, 1956.
3. V.V. Fedorov. *Theory of Optimal Experiments.* Academic Press, New York, 1972.
4. D.P. Hardin and E.B. Saff. Discretizing manifolds via minimum energy points. *Notices of the AMS*, 51(10):1186–1194, 2004.
5. M.E. Johnson, L.M. Moore, and D. Ylvisaker. Minimax and maximin distance designs. *Journal of Statistical Planning and Inference*, 26:131–148, 1990.
6. N.S. Landkof. *Foundations of Modern Potential Theory.* Springer, Berlin, 1972.
7. M.D. McKay, R.J. Beckman, and W.J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.
8. M.D. Morris and T.J. Mitchell. Exploratory designs for computational experiments. *Journal of Statistical Planning and Inference*, 43:381–402, 1995.
9. L. Pronzato and W.G. Müller. Design of computer experiments: space filling and beyond. *Statistics and Computing*, 22:681–701, 2012.
10. L. Pronzato and A. Pázman. *Design of Experiments in Nonlinear Models. Asymptotic Normality, Optimality Criteria and Small-Sample Properties.* Springer, LNS 212, New York, Heidelberg, 2013.
11. L. Pronzato, H.P. Wynn, and A. Zhigljavsky. Extended generalised variances, with applications. *Bernoulli (to appear)*, 2016. arXiv preprint arXiv:1411.6428.
12. E.B. Saff. Logarithmic potential theory with applications to approximation theory. *Surveys in Approximation Theory*, 5(14):165–200, 2010.
13. R.L. Schilling, R. Song, and Z. Vondracek. *Bernstein Functions: Theory and Applications.* de Gruyter, Berlin/Boston, 2012.
14. A.A. Zhigljavsky, H. Dette, and A. Pepelyshev. A new approach to optimal design for linear models with correlated observations. *Journal of the American Statistical Association*, 105(491):1093–1103, 2010.