Measures minimizing regularized dispersion

Luc Pronzato, Anatoly Zhigljavsky

Journal of Scientific Computing manuscript No.
(will be inserted by the editor)

Measures minimizing regularized dispersion

Luc Pronzato · **Anatoly Zhigljavsky**

Received: date / Accepted: date

Abstract We consider a continuous extension of a regularized version of the minimax, or dispersion, criterion widely used in space-filling design for computer experiments and quasi-Monte Carlo methods. We show that the criterion is convex for a certain range of the regularization parameter (depending on space dimension) and give a necessary and sufficient condition characterizing the optimal distribution of design points. Using results from potential theory, we investigate properties of optimal measures. The example of design in the unit ball is considered in details and some analytic results are presented. Using recent results and algorithms from experimental design theory, we show how to construct optimal measures numerically. They are often close to the uniform measure but do not coincide with it. The results suggest that designs minimizing the regularized dispersion for suitable values of the regularization parameter should have good space-filling properties. An algorithm is proposed for the construction of n-point designs.

Keywords dispersion · optimal design · space-filling design · potential theory

Mathematics Subject Classification (2000) 62K05 · 31C10 · 65D15

1 Introduction: dispersion and its regularization

Let $\mathscr{X}$ denote a compact subset of $\mathbb{R}^d$ with strictly positive d-dimensional Lebesgue measure and equal to the closure of its interior. Typical examples include the d-dimensional unit ball $\mathscr{B}_d(0,1) = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| \leq 1\}$ and convex polytopes. Throughout the paper, $\|.\|$ will denote the usual Euclidean norm, $\mathrm{vol}(\mathscr{X})$ the volume of $\mathscr{X}$, with $V_d = \mathrm{vol}[\mathscr{B}_d(0,1)] = \pi^{d/2}/\Gamma(d/2+1)$ the volume of $\mathscr{B}_d(0,1)$; μ will always denote the probability measure proportional to the Lebesgue measure on $\mathscr{X}$ (with therefore $\mu(\mathscr{X}) = 1$), but many of our results generalize to arbitrary probability measures on $\mathscr{X}$.

Consider an n-point set $X_n = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset \mathscr{X}$. The dispersion $\Phi(X_n)$ of X_n in $\mathscr{X}$ is defined by

$$\Phi(X_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_{i=1,\ldots,n} \|\mathbf{x} - \mathbf{x}_i\| ; \tag{1}$$

L. Pronzato (corresponding author)
Université Côte d'Azur, CNRS, I3S, France
Tel.: +33-4-92942708
Fax: +33-4-92942896
E-mail: Luc.Pronzato@cnrs.fr

A. Zhigljavsky
School of Mathematics, Cardiff University, UK
E-mail: ZhigljavskyAA@cardiff.ac.uk

see [14, Chap. 6]. It coincides with the Hausdorff distance between X_n and $\mathscr{X}$, with the minimax-distance criterion of [8], and is related to the covering radius of X_n, see [5]. As explained in [16] and many other references, n-point sets X_n with small dispersion are often desirable.

From classical properties of L_q and l_q norms, for any n-point set X_n we have $\Phi(X_n) = \lim_{q\to\infty} \Phi_q(X_n)$, where, for $q \neq 0$,

$$\Phi_q(X_n) = \Phi_q(X_n;\mu) = \left[\int_{\mathscr{X}} \left(\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{x}_i - \mathbf{x}\|^{-q}\right)^{-1} \mu(\mathrm{d}\mathbf{x})\right]^{1/q}, \tag{2}$$

which coincides with the coverage criterion suggested in [20] when $p = q$ in their notation; see also [17, Sect. 3.2].

Denote by Ξ the set of probability measures on $\mathscr{X}$ (called design measures) and let $\xi_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\mathbf{x}_i}$ denote the empirical design measure associated with X_n. Then we can also write $\Phi_q(X_n) = \phi_q(\xi_n)$, where, for any design measure $\xi \in \Xi$ we define the q-regularized dispersion by

$$\phi_q(\xi) = \phi_q(\xi;\mu) = \left[\int_{\mathscr{X}} \left(\int_{\mathscr{X}} \|\mathbf{x}-\mathbf{z}\|^{-q}\,\xi(\mathrm{d}\mathbf{z})\right)^{-1} \mu(\mathrm{d}\mathbf{x})\right]^{1/q}, \; q \neq 0. \tag{3}$$

Straightforward calculation yields the following extension of $\phi_q(\xi)$ for $q = 0$:

$$\phi_0(\xi) = \exp\int_{\mathscr{X}}\int_{\mathscr{X}} \log\|\mathbf{x}-\mathbf{z}\|\,\xi(\mathrm{d}\mathbf{z})\,\mu(\mathrm{d}\mathbf{x}).$$

We shall investigate the properties of regularized dispersions $\phi_q(\cdot)$ and characterize optimal measures that minimize them. Our results suggest that designs obtained by minimizing $\Phi_q(\cdot)$ defined in (2) with $q > 0$ between $d-2$ and d should have good space-filling properties.

The paper is organized as follows. Properties of regularized dispersion are considered in Section 2. The cases $q < 0$ and $q \geq d$ are of limited interest and will only be briefly considered. Convexity of the functional $\phi_q^q(\cdot)$ for $q \geq 0$ is proved in Section 2.2 and a necessary and sufficient condition for design optimality is derived in Section 2.3 for the case $0 < q < d$. In Section 3, we show that optimal measures can be singular when $0 < q \leq d-2$ but do not contain any atoms in the interior of $\mathscr{X}$ when $\max\{0, d-2\} < q < d$. A comparison with the behaviour of minimum-energy measures, which maximize a regularized maximin-distance criterion, is discussed in Section 3.4. The algorithmic construction of optimal measures and numerical examples are presented in Section 4, and the generation of n-point sets with small dispersion is considered in Section 5. Section 6 draws some conclusions.

2 Properties of regularized dispersion

Note that for any $q \in \mathbb{R}$ and any measure $\xi \in \Xi$, we have $0 \leq \phi_q(\xi) \leq \mathrm{diam}(\mathscr{X})$, where $\mathrm{diam}(\mathscr{X}) = \max_{\mathbf{x},\mathbf{z}\in\mathscr{X}} \|\mathbf{x}-\mathbf{z}\|$ is the diameter of $\mathscr{X}$. When $\xi = \delta_{\mathbf{z}}$, the delta measure at some $\mathbf{z} \in \mathscr{X}$, then

$$\phi_q(\delta_{\mathbf{z}}) = \left[\int_{\mathscr{X}} \|\mathbf{z}-\mathbf{x}\|^q\,\mu(\mathrm{d}\mathbf{x})\right]^{1/q}. \tag{4}$$

By taking the derivative with respect to q and using Jensen's inequality for the function $t \to t\log t$ we can observe that $\phi_q(\delta_{\mathbf{z}})$ is an increasing function of $q \in \mathbb{R}$. Also, $\phi_q(\delta_{\mathbf{z}}) > 0$ for $q > -d$ and $\phi_q(\delta_{\mathbf{z}}) = 0$ for $q \leq -d$.

From the convexity of $f(t) = 1/t$ and Jensen's inequality, for any real q and any $\xi \in \Xi$, we have $\int_{\mathscr{X}} \|\mathbf{z}-\mathbf{x}\|^q\,\xi(\mathrm{d}\mathbf{z}) \geq [\int_{\mathscr{X}} \|\mathbf{z}-\mathbf{x}\|^{-q}\,\xi(\mathrm{d}\mathbf{z})]^{-1}$ for any $\mathbf{x} \in \mathscr{X}$. This implies

$$\begin{aligned}\max_{\mathbf{z}\in\mathscr{X}} \int_{\mathscr{X}} \|\mathbf{z}-\mathbf{x}\|^q\,\mu(\mathrm{d}\mathbf{x}) &\geq \int_{\mathscr{X}}\int_{\mathscr{X}} \|\mathbf{z}-\mathbf{x}\|^q\,\mu(\mathrm{d}\mathbf{x})\,\xi(\mathrm{d}\mathbf{z}) \\ &\geq \int_{\mathscr{X}} \left[\int_{\mathscr{X}} \|\mathbf{z}-\mathbf{x}\|^{-q}\,\xi(\mathrm{d}\mathbf{z})\right]^{-1} \mu(\mathrm{d}\mathbf{x}) = \phi_q^q(\xi).\end{aligned}$$

Therefore, when $q < 0$,

$$\phi_q(\xi) \geq \left[\int_{\mathscr{X}} \|\mathbf{z}^* - \mathbf{x}\|^q \, \mu(\mathrm{d}\mathbf{x})\right]^{1/q} = \phi_q(\delta_{\mathbf{z}^*}),$$

where $\mathbf{z}^*$ is any point such that

$$\int_{\mathscr{X}} \|\mathbf{z}^* - \mathbf{x}\|^q \, \mu(\mathrm{d}\mathbf{x}) = \max_{\mathbf{z}\in\mathscr{X}} \int_{\mathscr{X}} \|\mathbf{z} - \mathbf{x}\|^q \, \mu(\mathrm{d}\mathbf{x}). \tag{5}$$

When $q < 0$, ϕ_q-optimal measures minimizing $\phi_q(\cdot)$ are thus delta measures concentrated at any such $\mathbf{z}^*$. In particular, when $\mathscr{X}$ contains its center of symmetry $\mathbf{c}$, then $\mathbf{z}^* = \mathbf{c}$. We shall only consider the case $q \geq 0$ in the rest of the paper.

2.1 The case $q \geq d$

Suppose that the measure ξ satisfies $\xi[\mathscr{B}_d(\mathbf{z}_0, r)] = 0$ for some ball $\mathscr{B}_d(\mathbf{z}_0, r) \subset \mathscr{X}$ having center $\mathbf{z}_0$ and radius $r > 0$. Then, $\|\mathbf{z} - \mathbf{x}\| > r/2$ for any $\mathbf{x} \in \mathscr{B}_d(\mathbf{z}_0, r/2)$ and any $\mathbf{z}$ in the support of ξ, so that

$$\phi_q^q(\xi) \geq \int_{\mathscr{B}_d(\mathbf{z}_0, r/2)} (r/2)^q \, \mu(\mathrm{d}\mathbf{x}) = \frac{V_d}{\mathrm{vol}(\mathscr{X})} (r/2)^{d+q} > 0,$$

which shows in particular that $\phi_q(\xi) > 0$ for any $q \geq 0$ when ξ is a discrete measure.

On the other hand, when $q \geq d$ then $\phi_q(\xi) = 0$ for any $\xi \in \Xi$ equivalent to the Lebesgue measure on $\mathscr{X}$. Any such measure is therefore ϕ_q-optimal. Note that, similarly to Theorem 1, we can prove that the functionals $\phi_q^q(\cdot)$ are convex for $q \geq d$.

2.2 Properties of $\phi_q(\cdot)$ when $0 \leq q < d$

For this range of q, $\phi_q(\xi) > 0$ for any $\xi \in \Xi$. This follows from the fact that Riesz potentials

$$P_{\nu,q}(\mathbf{x}) = \int_{\mathscr{X}} \|\mathbf{x} - \mathbf{z}\|^{-q} \, \nu(\mathrm{d}\mathbf{z}) \tag{6}$$

are such that, for any positive measure ν on $\mathscr{X}$, $P_{\nu,q}(\mathbf{x})$ is finite for μ-almost all $\mathbf{x}$ when $q < d$; see [9, p. 61]. The key property for the determination of ϕ_q-optimal measures is convexity of the functionals $\phi_0(\cdot)$ and $\phi_q^q(\cdot)$ when $q > 0$.

Theorem 1 *The functional $\phi_0(\cdot)$ is convex on Ξ and the functionals $\phi_q^q(\cdot)$ are strictly convex for $0 < q < d$.*

Proof The convexity of $\phi_0(\cdot)$ directly follows from the convexity of the exponential function. Suppose now that $q > 0$ and consider $\xi_\alpha = (1-\alpha)\xi_0 + \alpha\xi_1$ for ξ_0 and ξ_1 two arbitrary measures in Ξ and $\alpha \in [0,1]$. Direct calculation gives

$$\frac{\mathrm{d}\phi_q^q(\xi_\alpha)}{\mathrm{d}\alpha} = -\int_{\mathscr{X}} \left\{ \left(\int_{\mathscr{X}} \|\mathbf{x} - \mathbf{z}\|^{-q} \, \xi_\alpha(\mathrm{d}\mathbf{z}) \right)^{-2} \int_{\mathscr{X}} \|\mathbf{x} - \mathbf{z}\|^{-q} \, (\xi_1 - \xi_0)(\mathrm{d}\mathbf{z}) \right\} \mu(\mathrm{d}\mathbf{x}) \tag{7}$$

and

$$\left.\frac{\mathrm{d}^2\phi_q^q(\xi_\alpha)}{\mathrm{d}\alpha^2}\right|_{\alpha=0} = 2\int_{\mathscr{X}} \left\{ \left(\int_{\mathscr{X}} \|\mathbf{x} - \mathbf{z}\|^{-q} \, \xi_0(\mathrm{d}\mathbf{z}) \right)^{-3} \right.$$
$$\left. \times \left(\int_{\mathscr{X}} \|\mathbf{x} - \mathbf{z}\|^{-q} \, (\xi_1 - \xi_0)(\mathrm{d}\mathbf{z}) \right)^2 \right\} \mu(\mathrm{d}\mathbf{x}).$$

Since this second-order derivative is non-negative, the functional $\phi_q^q(\cdot)$ is convex. Let us show that it is strictly convex. Since $q < d$, $P_{\xi_0,q}(\mathbf{x}) = \int_{\mathscr{X}} \|\mathbf{x}-\mathbf{z}\|^{-q}\,\xi_0(\mathrm{d}\mathbf{z})$ is finite for μ-almost all $\mathbf{x}$; see [9, p. 61]. Moreover, $P_{\xi_1,q}(\mathbf{x}) = P_{\xi_0,q}(\mathbf{x})$ implies that $\xi_0 = \xi_1$; see [9, p. 74]. This yields $\mathrm{d}^2\phi_q^q(\xi_\alpha)/\mathrm{d}\alpha^2\big|_{\alpha=0} > 0$ for $\xi_1 \neq \xi_0$. □

Note that $\phi_0(\cdot)$ is clearly not strictly convex. On the other hand, the minimum of $\phi_0(\xi)$ is attained when ξ is the delta measure at any $\mathbf{z}^* = \arg\min_{\mathbf{z}\in\mathscr{X}} \int_{\mathscr{X}} \log\|\mathbf{z}-\mathbf{x}\|\,\mu(\mathrm{d}\mathbf{x})$. As in the case $q < 0$, if $\mathbf{c}$ is a center of symmetry for $\mathscr{X}$ and $\mathbf{c} \in \mathscr{X}$, then $\mathbf{z}^* = \mathbf{c}$. Together with (5), this solves the problem of determination of ϕ_q-optimal measures for $q \leq 0$, for any set $\mathscr{X}$. The case $q > d$ has been solved in Section 2.1. Hence, only the case $0 < q < d$ is considered in the rest of the paper (except in Section 5, where values $q \geq d$ are used to construct n-point sets).

2.3 Necessary and sufficient condition for optimality

From Theorem 1, when $0 < q < d$ there exists a unique design measure $\xi^{q,*}$ that minimizes $\phi_q(\xi)$ with respect to $\xi \in \Xi$. The convexity of $\phi_q^q(\cdot)$ yields a necessary and sufficient condition for the optimality of $\xi^{q,*}$, as expressed by the following theorem. We denote by $F_q(\xi,\mathbf{y})$ the directional derivative of $\phi_q^q(\cdot)$ at ξ in the direction of the delta measure at $\mathbf{y}$,

$$F_q(\xi,\mathbf{y}) = \frac{\mathrm{d}\phi_q^q[(1-\alpha)\xi + \alpha\delta_{\mathbf{y}}]}{\mathrm{d}\alpha}\bigg|_{\alpha=0}. \tag{8}$$

Theorem 2 *The measure $\xi^{q,*}$ is ϕ_q-optimal ($0 < q < d$) if and only if*

$$\forall \mathbf{y} \in \mathscr{X}\,,\; d(\xi^{q,*},\mathbf{y}) \leq \phi_q^q(\xi^{q,*})\,, \tag{9}$$

where, for any $\xi \in \Xi$ and $\mathbf{y} \in \mathscr{X}$,

$$d(\xi,\mathbf{y}) = \phi_q^q(\xi) - F_q(\xi,\mathbf{y}) = \int_{\mathscr{X}} \left\{ \|\mathbf{y}-\mathbf{x}\|^{-q} \left(\int_{\mathscr{X}} \|\mathbf{x}-\mathbf{z}\|^{-q}\,\xi(\mathrm{d}\mathbf{z}) \right)^{-2} \right\} \mu(\mathrm{d}\mathbf{x})\,. \tag{10}$$

Moreover, $\xi^{q,*}$ is unique and $d(\xi^{q,*},\mathbf{y}) = \phi_q^q(\xi^{q,*})$ for $\xi^{q,*}$-almost all $\mathbf{y} \in \mathscr{X}$.

Proof The proof of (9) follows from the expression (7) using the property that $\xi^{q,*} \in \Xi$ minimizes $\phi_q(\xi)$ if and only if $\mathrm{d}\phi_q^q(\xi_\alpha)/\mathrm{d}\alpha\big|_{\alpha=0} \geq 0$ for any $\xi_1 \in \Xi$ when we write $\xi_\alpha = (1-\alpha)\xi^{q,*} + \alpha\xi_1$. Noticing that $\int_{\mathscr{X}} d(\xi,\mathbf{y})\,\xi(\mathrm{d}\mathbf{y}) = \phi_q^q(\xi)$ for any $\xi \in \Xi$, we obtain that $d(\xi^{q,*},\mathbf{y}) = \phi_q^q(\xi^{q,*})$, $\xi^{q,*}$-almost everywhere. The uniqueness of $\xi^{q,*}$ follows from the strict convexity of $\phi_q^q(\cdot)$ established in Theorem 1. □

Remark 1 *Consider the measure ν_ξ defined by*

$$\nu_\xi(\mathrm{d}\mathbf{x}) = \left[\int_{\mathscr{X}} \|\mathbf{x}-\mathbf{z}\|^{-q}\,\xi(\mathrm{d}\mathbf{z}) \right]^{-2} \mu(\mathrm{d}\mathbf{x})\,. \tag{11}$$

We can then write $d(\xi,\mathbf{y}) = P_{\nu_\xi,q}(\mathbf{y})$, where $P_{\nu,q}(\cdot)$ is the Riesz potential for the measure ν, as defined in (6)*.*

In the next section we show how to use Theorem 2 to characterize optimal design measures.

3 Properties of optimal design measures

We consider the two cases $0 < q \leq d-2$ and $\max\{0, d-2\} < q < d$ separately.

3.1 $0 < q \leq d-2$

Direct calculations give the following expressions for the first and second-order derivatives of the function $d(\xi,\mathbf{y})$:

$$\frac{\partial d(\xi,\mathbf{y})}{\partial \mathbf{y}} = -q\int_{\mathscr{X}} (\mathbf{y}-\mathbf{x})\|\mathbf{y}-\mathbf{x}\|^{-q-2}\,\nu_\xi(\mathrm{d}\mathbf{x})\,,$$
$$\frac{\partial^2 d(\xi,\mathbf{y})}{\partial \mathbf{y}\partial \mathbf{y}^\top} = q\int_{\mathscr{X}} \|\mathbf{y}-\mathbf{x}\|^{-q-4}\left[(q+2)(\mathbf{y}-\mathbf{x})(\mathbf{y}-\mathbf{x})^\top - \|\mathbf{y}-\mathbf{x}\|^2\mathbf{I}_d\right]\,\nu_\xi(\mathrm{d}\mathbf{x})\,,$$

where $\nu_\xi(\mathrm{d}\mathbf{x})$ is the measure defined by (11) and $\mathbf{I}_d$ denotes the $d\times d$ identity matrix. One can readily notice that for all $\mathbf{y}\in\mathscr{X}$ the Laplacian $\Delta d(\xi,\mathbf{y})$ satisfies

$$\Delta d(\xi,\mathbf{y}) = \mathrm{trace}\left[\frac{\partial^2 d(\xi,\mathbf{y})}{\partial y\partial \mathbf{y}^\top}\right] = q(q+2-d)\int_{\mathscr{X}} \|\mathbf{y}-\mathbf{x}\|^{-q-2}\,\nu_\xi(\mathrm{d}\mathbf{x})\,, \tag{12}$$

which is finite and non-positive for $0<q\leq d-2$. This corresponds to the fact that the Riesz potentials $P_{\nu,q}(\cdot)$ of Remark 1 are superharmonic on $\mathbb{R}^d$ for this range of q; see [9, p. 66]. The minimum principle for superharmonic functions implies that $d(\xi^{q,*},\mathbf{y})$ cannot reach its minimum in the interior of $\mathscr{X}$; therefore, either $\xi^{q,*}$ is fully supported away from the boundary of $\mathscr{X}$ or $d(\xi^{q,*},\mathbf{y})$ is constant over $\mathscr{X}$. Typically, the later would mean that $\xi^{q,*}$ is equivalent to μ.

Using the property of superharmonicity of $d(\xi^{q,*},\cdot)$, we can construct examples where the ϕ_q-optimal measure is singular.

Example 1
Take $\mathscr{X}=\mathscr{B}_d(0,1)$ and $\xi=\delta_0$, the delta measure at the centre of the ball. The function $d(\delta_0,\mathbf{y})=P_{\nu_{\delta_0},q}(\mathbf{y})$ only depends on $\|\mathbf{y}\|$, and $d(\delta_0,\mathbf{y})=\int P_{\nu_{\delta_0},q}(\mathbf{z})\,\mu_{\mathscr{S}_d(0,\|\mathbf{y}\|)}(\mathrm{d}\mathbf{z})$, where $\mu_{\mathscr{S}_d(0,r)}$ is the normalized surface measure on the sphere $\mathscr{S}_d(0,r)$ of radius r centered at the origin. Since $P_{\nu,q}(\cdot)$ is superharmonic on $\mathbb{R}^d$,

$$\int P_{\nu_{\delta_0},q}(\mathbf{z})\,\mu_{\mathscr{S}_d(0,\|\mathbf{y}\|)}(\mathrm{d}\mathbf{z}) \leq P_{\nu_{\delta_0},q}(0)\,;$$

see [9, p. 52]. Direct calculation shows that $P_{\nu_{\delta_0},q}(0)=d(\delta_0,0)=\phi_q^q(\delta_0)$. Therefore, $d(\delta_0,\mathbf{y})\leq \phi_q^q(\delta_0)$ for all $\mathbf{y}$ in $\mathscr{X}$, and Theorem 2 implies that the design measure δ_0 is ϕ_q-optimal.

On the other hand, we show below that δ_0 is no longer ϕ_q-optimal when the ball is deformed into an ellipsoid. Consider $\mathscr{X}=\{\mathbf{x}\in\mathbb{R}^d : \mathbf{x}^\top\mathbf{A}\mathbf{x}\leq 1\}$ with $\mathbf{A}=\mathrm{diag}\{a^2,1,\ldots,1\}$. For $\mathbf{y}=(r,0,\ldots,0)$, we get

$$d(\delta_0,\mathbf{y}) = P_{\nu_{\delta_0},q}(\mathbf{y}) = \int_{\mathscr{X}} \frac{\|\mathbf{x}\|^{2q}}{[(x_1-r)^2+\rho^2]^{q/2}}\,\mu(\mathrm{d}\mathbf{x})$$
$$= \frac{1}{a}\int_{\mathscr{B}_d(0,1)} \frac{[z_1^2/a^2+\rho^2]^q}{[(z_1/a-r)^2+\rho^2]^{q/2}}\,\mu(\mathrm{d}\mathbf{z})\,,$$

where $\rho^2=\sum_{i\geq 2}x_i^2$ and $\mathbf{z}=(ax_1,x_2,\ldots,x_d)$. The marginal density of z_1 for $\mathbf{z}$ uniformly distributed in $\mathscr{B}_d(0,1)$ is

$$\varphi_d(z_1) = \frac{V_{d-1}}{V_d}(1-z_1^2)^{(d-1)/2}\,,\; z_1\in[-1,1]\,,$$

and the conditional density of ρ given z_1 is

$$\varphi_d(\rho|z_1) = \frac{(d-1)\rho^{d-2}}{(1-z_1^2)^{(d-1)/2}}\,,\; \rho\in[0,\sqrt{1-z_1^2}]\,.$$

This gives

$$d(\delta_0,\mathbf{y}) = \frac{(d-1)V_{d-1}}{2aV_d} \int_{-1}^{1} \int_{0}^{1-z_1^2} \frac{[z_1^2/a^2+s]^q}{[(z_1/a-r)^2+s]^{q/2}}\, s^{(d-3)/2}\, \mathrm{d}s\, \mathrm{d}z_1\,,$$

where we made the change of variable $s=\rho^2$. The function $d(\delta_0,y)$, for any given r, can thus be made arbitrarily large by taking a small enough.

Example 2

Take $q=2$ and $\mathscr{X}$ the d-dimensional cube $[-1/2,1/2]^d$ with $d=4$. Figure 1 shows the values of $d(\delta_0,\mathbf{y})$, obtained by numerical integration, as a function of $\|\mathbf{y}\|$ when $\mathbf{y}$ moves along a principal axis (dashed line) and along the main diagonal of $\mathscr{X}$ (solid line). The figure shows that, although the delta measure δ_0 is ϕ_q-optimal for $\mathscr{X}=\mathscr{B}_4(0,1)$ with $q=2$, it is not optimal for the cube.

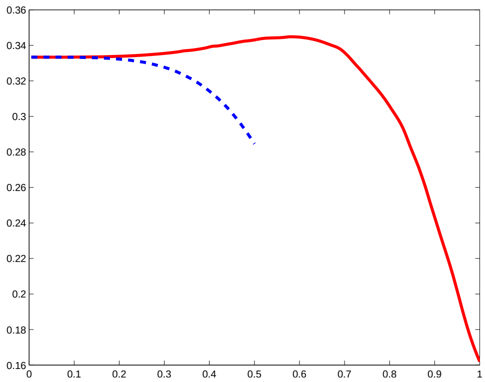

Fig. 1 $d(\delta_0,\mathbf{y})$ as a function of $\|\mathbf{y}\|$ when $\mathbf{y}$ moves along a principal axis (dashed line) or along the main diagonal (solid line) of the cube $[-1/2,1/2]^4$, $q=d-2=2$.

3.2 $\max\{0,d-2\}<q<d$

From the expression (12) of the second-order derivative of $d(\xi,\mathbf{y})$, we can deduce the following property concerning the presence of atoms in a ϕ_q-optimal design measure.

Theorem 3 *For any $q\in(\max\{0,d-2\},d)$, the ϕ_q-optimal design measure $\xi^{q,*}$ does not contain atoms in the interior of $\mathscr{X}$.*

Proof Assume ξ is ϕ_q-optimal and has an atom at some $\mathbf{z}$ in the interior of $\mathscr{X}$. This implies that $\partial d(\xi,\mathbf{y})/\partial\mathbf{y}\big|_{\mathbf{y}=\mathbf{z}}=0$. Let us write ξ as $\xi=w\delta_{\mathbf{z}}+(1-w)\xi'$, where $w>0$ and ξ' is any probability measure on $\mathscr{X}$. Then, (12) gives

$$\mathrm{trace}\left[\left.\frac{\partial^2 d(\xi,\mathbf{y})}{\partial\mathbf{y}\partial\mathbf{y}^\top}\right|_{\mathbf{y}=\mathbf{z}}\right] = q(q+2-d)\int_{\mathscr{X}} \|\mathbf{z}-\mathbf{x}\|^{-q-2} \times\left(w\|\mathbf{x}-\mathbf{z}\|^{-q}+(1-w)\int_{\mathscr{X}}\|\mathbf{x}-\mathbf{y}\|^{-q}\,\xi'(\mathrm{d}\mathbf{y})\right)^{-2}\mu(\mathrm{d}\mathbf{x})\,,$$

which is strictly positive, bounded by $[q(q+2-d)/w^2]\int_{\mathscr{X}}\|\mathbf{z}-\mathbf{x}\|^{q-2}\,\mu(\mathrm{d}\mathbf{x})<\infty$ and thus well defined. Therefore, $\max\{0,d-2\}<q<d$ implies that

$$\mathbf{e}_i^\top\,\frac{\partial^2 d(\xi,\mathbf{y})}{\partial\mathbf{y}\partial\mathbf{y}^\top}\,\mathbf{e}_i>0$$

for at least one basis vector $\mathbf{e}_i$. Since $\partial d(\xi,\mathbf{y})/\partial \mathbf{y}\big|_{\mathbf{y}=\mathbf{z}} = 0$, it implies that there exists a $\mathbf{y}$ in the neighborhood of $\mathbf{z}$ such that $d(\xi,\mathbf{y}) > d(\xi,\mathbf{z})$, which contradicts the optimality condition for ξ in Theorem 2. □

Examples of Section 4.4 will present situations where $\xi^{q,*}$ has a density with respect to the Lebesgue measure.

3.3 Spherically symmetric design measures in the unit ball

Assume that $\mathscr{X} = \mathscr{B}_d(0,1)$ and $0 < q < d$. By symmetry, ϕ_q-optimal mesures are spherically symmetric, and we can thus restrict our attention to such measures. Also, the integral $\int_{\mathscr{X}} \|\mathbf{x}-\mathbf{z}\|^{-q}\,\xi(\mathrm{d}\mathbf{z})$ in (3) only depends on $\mathbf{x}$ through $\rho = \|\mathbf{x}\|$, with ρ having the density $d\rho^{d-1}$, $\rho \in [0,1]$. We can thus write

$$\phi_q^q(\xi) = d \int_0^1 J_{d,q}^{-1}(\rho)\,\rho^{d-1}\,\mathrm{d}\rho\,, \tag{13}$$

where $J_{d,q}(\rho) = \int_{\mathscr{X}} \|\mathbf{x}_\rho - \mathbf{z}\|^{-q}\,\xi(\mathrm{d}\mathbf{z})$, with $\|\mathbf{x}_\rho\| = \rho$. Let ω_ξ denote the probability measure of $R = \|\mathbf{z}\|$ when $\mathbf{z} \sim \xi$, then

$$J_{d,q}(\rho) = \int_0^1 I_{d,q}(R,\rho)\,\omega_\xi(\mathrm{d}R)\,, \tag{14}$$

where

$$I_{d,q}(R,\rho) = \int_{\mathscr{S}_d(0,R)} \|\mathbf{x}_\rho - \mathbf{z}\|^{-q}\,\xi_R(\mathrm{d}\mathbf{z}) \tag{15}$$

and $\xi_R(\mathrm{d}\mathbf{z})$ denotes the probability measure of $\mathbf{z} \sim \xi$ conditional on $\|\mathbf{z}\| = R$. Take $\mathbf{x}_\rho = (\rho,0,\ldots,0)$, decompose $\|\mathbf{x}_\rho - \mathbf{z}\|$ as $\|\mathbf{x}_\rho - \mathbf{z}\| = [(z_1-\rho)^2 + \sum_{i=2}^d z_i^2]^{1/2} = [(z_1-\rho)^2 + R^2 - z_1^2]^{1/2}$, and denote

$$\psi_d(t|R) = \frac{(d-1)V_{d-1}}{dV_d}\,\frac{(R^2-t^2)^{(d-3)/2}}{R^{d-2}}\,,\ t \in [-R,R],$$

the density of the first component $t = z_1$ of $\mathbf{z} = (z_1,\ldots,z_d)$ uniformly distributed on the sphere $\mathscr{S}_d(0,R)$. We can then write

$$I_{d,q}(R,\rho) = \int_{-R}^{R} [(t-\rho)^2 + R^2 - t^2]^{-q/2}\,\psi_d(t|R)\,\mathrm{d}t\,. \tag{16}$$

Using the expression (16), we can express the integral (15) in terms of the hypergeometric function ${}_2\mathrm{F}_1(\cdot,\cdot;\cdot;\cdot)$ as follows:

$$I_{d,q}(R,\rho) = \left(R^2+\rho^2\right)^{-q/2}\,{}_2\mathrm{F}_1\left(\frac{q}{4},\frac{q+2}{4};\frac{d}{2};\frac{4\rho^2R^2}{(R^2+\rho^2)^2}\right),$$

which reveals the full symmetry between R and ρ. Explicit expressions of $I_{d,q}(R,\rho)$ for particular values of d and q are given in Appendix, together with values of $J_{d,q}(\rho)$, see (14), when $\mathbf{z}$ is uniformly distributed in $\mathscr{B}_d(0,1)$ (so that $\omega_\xi(\mathrm{d}R) = dR^{d-1}\,\mathrm{d}R$).

3.4 Comparison with regularized maximin-distance optimal design

The maximin-distance criterion for a given n-point set $X_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathscr{X}$ is

$$\Psi(X_n) = \min_{i \neq j} \|\mathbf{x}_i - \mathbf{x}_j\| \,; \tag{17}$$

see [8]. Although the criterion $\Psi(\cdot)$ is less intuitively appealing than $\Phi(\cdot)$ defined in (1), designs that maximize $\Psi(X_n)$ for a given n are also often considered in computer experiments; see [17] and [16, Sect. 2]. The regularized version $\Psi_q(X_n)$ of $\Psi(X_n)$ is similar to $\Phi_q(X_n)$ given by (2):

$$\Psi_q(X_n) = \left[\frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \|\mathbf{x}_i - \mathbf{x}_j\|^{-q} \right]^{-1/q} \,, \; q > 0 \,.$$

Maximization of $\Psi_q(X_n)$ is equivalent to the determination of n Fekete points $\mathbf{x}_i$ that minimize the q-energy $\sum_{1 \leq i < j \leq n} \|\mathbf{x}_i - \mathbf{x}_j\|^{-q}$; see for example [7], [19]. The continuous version of this energy is the functional

$$E_q(\zeta) = \int_{\mathscr{X}} \int_{\mathscr{X}} \|\mathbf{z} - \mathbf{t}\|^{-q} \, \zeta(\mathrm{d}\mathbf{t}) \, \zeta(\mathrm{d}\mathbf{z}) \,.$$

The determination of minimum-energy probability measures $\zeta^{q,*}$ (called q-equilibrium measures) is one of the main topics in potential theory; see [9]. A typical result is as follows. For $d-2 < q < d$, $\zeta^{q,*}$ is uniquely defined and is equivalent to the Lebesgue measure on $\mathscr{X}$, showing some similarity with the situation in Section 3.2. For $0 < q \leq d-2$, $\zeta^{q,*}$ is concentrated on the boundary of $\mathscr{X}$, a configuration opposite to that encountered in Section 3.1, see Example 1.

4 Algorithmic constructions of discrete approximations to optimal measures

4.1 An A-optimal design problem

Consider a discrete approximation of μ, given by $\mu_N = (1/N) \sum_{i=1}^N \delta_{\mathbf{x}_i}$, where the $\mathbf{x}_i$ are taken from a regular grid or correspond to the first N points of a low-discrepancy sequence in $\mathscr{X}$. Denote by $\mathscr{M}(\mathbf{z})$ the $N \times N$ matrix $\mathscr{M}(\mathbf{z}) = \mathrm{diag}\{N \, \|\mathbf{z} - \mathbf{x}_i\|^{-q}, \, i = 1, \ldots, N\}$. Then,

$$\phi_q^q(\xi; \mu_N) = \mathrm{trace}[\mathbf{M}^{-1}(\xi)] \,,$$

where $\mathbf{M}(\xi) = \int_{\mathscr{X}} \mathscr{M}(\mathbf{z}) \, \xi(\mathrm{d}\mathbf{z})$ can be considered as an information matrix for the measure ξ (approximate design theory), with $\mathscr{M}(\mathbf{z})$ the elementary information matrix for $\mathbf{z}$. This indicates that the determination of a $\phi_q(\cdot; \mu_N)$-optimal measure corresponds to an A-optimal design problem. From Caratheodory theorem, for any $\xi \in \Xi$, there always exists a discrete measure $\xi_m \in \Xi$ supported on $m \leq N+1$ points in $\mathscr{X}$ and such that $\mathbf{M}(\xi) = \mathbf{M}(\xi_m)$, and therefore $\phi_q(\xi; \mu_N) = \phi_q(\xi_m; \mu_N)$. On the other hand, for any given $\xi \in \Xi$, if μ_N tends to μ (weak convergence of probability measures) then $\phi_q(\xi; \mu_N)$ tends to $\phi_q(\xi) = \phi_q(\xi; \mu)$. When μ_N corresponds to a low discrepancy sequence, Koksma-Hlawka inequality, see [14, p. 20], gives the error bound $|\phi_q(\xi; \mu_N) - \phi_q(\xi)| < C \, (\log N)^d / N$ for some constant C.

4.2 Algorithms

From Section 4.1, when considering a discrete approximation of μ, given by $\mu_N = (1/N) \sum_{i=1}^N \delta_{\mathbf{x}_i}$, the construction of a ϕ_q-optimal design measure corresponds to an A-optimal design problem, for which several optimization algorithms are available; see, e.g., [18, Chap. 9].

Consider the approximation of ξ by a discrete measure, $\xi_m = \sum_{j=1}^m w_j \delta_{\mathbf{z}_j}$, where the weights w_i are positive and sum to one. The $\mathbf{z}_j$ are fixed and such that $\|\mathbf{z}_j - \mathbf{x}_i\| > 0$ for any i, j. For instance, one may take the $\mathbf{x}_i$ as the first N points of a low-discrepancy sequence in $\mathscr{X}$, and the $\mathbf{z}_j$ as the next m points of the same sequence.

The measure ξ is then characterized by the m-dimensional vector of weights $\mathbf{w}$, with components w_j. The determination of the optimal vector of weights $\mathbf{w}^*$, unique from the strict convexity of $\phi_q(\cdot)$, forms a convex problem. However, the high dimensionality of $\mathbf{w}$ makes the direct use of nonlinear programming methods rather inefficient, and two methods from the experimental design literature are suggested below.

4.2.1 Multiplicative algorithm

Let $\xi_m^{(k)}$ denote the design measure at iteration k, with weight $w_j^{(k)}$ assigned to $\mathbf{z}_j$, $j = 1,\ldots,m$. Suppose that $w_j^{(0)} > 0$ for all j. In [22], the algorithm defined by

$$w_j^{(k+1)} = w_j^{(k)} \frac{d^\lambda(\xi_m^{(k)},\mathbf{z}_j;\mu_N)}{\sum_{\ell=1}^m d^\lambda(\xi_m^{(k)},\mathbf{z}_\ell;\mu_N)} \tag{18}$$

is proved to converge monotonically to the optimal measure for $\phi_q(\cdot;\mu_N)$ supported on the $\mathbf{z}_j$, for any $\lambda \in (0,1)$.

4.2.2 Vertex-exchange algorithm

We consider the method proposed in [3,4] and [10,11], and denote

$$j_+ = \arg\max_j d(\xi_m^{(k)},\mathbf{z}_j;\mu_N) \;\text{ and }\; j_- = \arg\min_j d(\xi_m^{(k)},\mathbf{z}_j;\mu_N)$$

(take any of them in case there are several solutions), with $\xi_m^{(k)}$ the design measure at iteration k. The updating rule is given by

$$w_{j_+}^{(k+1)} = w_{j_+}^{(k)} + \alpha^k\,,\; w_{j_-}^{(k+1)} = w_{j_-}^{(k)} - \alpha^k \text{ and } w_j^{(k+1)} = w_j^{(k)} \text{ for all j } \neq j_+, j_-\,,$$

with $\alpha^k \in [0, w_{j_-}^{(k)}]$. The value of α^k can be chosen by minimizing $\phi_q(\xi_m^{(k+1)};\mu_N)$, but an Armijo type line-search [1] is generally very efficient. Note that the point $\mathbf{z}_{j_-}$ is removed from the support of $\xi_m^{(k)}$ when $\alpha^k = w_{j_-}^{(k)}$; a more general procedure for reducing the support of $\xi_m^{(k)}$ is presented in the next section.

4.2.3 Support reduction

Using the equivalence between $\phi_q(\cdot;\mu_N)$- and A-optimal designs, and the results in [15], we can derive an inequality that the support points of the optimal design must satisfy. In many situations, this gives a simple characterization of an outer set of this support. Removing points that cannot be in the support of the optimal design also reduces the amount of computations in the algorithms above.

From [15, Theorem 2], any $\mathbf{z}$ such that

$$\text{trace}[\mathscr{M}(\mathbf{z})\mathbf{M}^{-2}(\xi)] < \kappa^2 B(t,\varepsilon)$$

cannot be in the support of the optimal measure for $\phi_q(\cdot;\mu_N)$. Here, $B(t,\varepsilon) = t(1+\varepsilon/t)^{-1}$, with $t = \text{trace}[\mathbf{M}^{-1}(\xi)]$ and $\varepsilon = \max_{\mathbf{z}\in\mathscr{X}} \text{trace}[\mathscr{M}(\mathbf{z})\mathbf{M}^{-2}(\xi)] - t$, and κ is the unique solution in $(\sqrt{\alpha},1]$ of the equation in θ

$$\frac{\alpha}{\theta^2} + \frac{(1-\alpha)^3}{(1+\varepsilon/t-\alpha\theta)^2} = 1\,,$$

where $\alpha = \lambda_{\min}[\mathbf{M}^{-1}(\xi)]/t$.

4.3 Algorithmic constructions in the unit ball

The constructions described in Sections 4.1 and 4.2 suppose that one is able to generate dense enough sequences in $\mathscr{X}$, and are thus restricted to very small dimensions d (typically, $d = 1$ or 2). On the other hand, when $\mathscr{X}$ is the unit ball $\mathscr{B}_d(0,1)$, using the representation given by (13) and (14) for $\phi_q(\xi)$ when ξ is spherically symmetric, together with the expression in Appendix for $I_{d,q}(R,\rho)$, we only need to use one-dimensional sequences for any $d > 1$. The construction is as follows.

First, we generate a low discrepancy point-set $u_1,\ldots,u_N$ in $[0,1]$, and transform it into $\rho_i = u_i^{1/d}$, so that the ρ_i have the same distribution as the radii of points uniformly distributed in $\mathscr{B}_d(0,1)$. Each ρ_i receives the same weight $1/N$ to approximate the uniform measure μ in $\mathscr{B}_d(0,1)$. We then proceed similarly for the radii R_j, $j = 1,\ldots,m$, each of them receiving some weight w_j, with $w_j = w_j^{(0)} > 0$, for instance $w_j^{(0)} = 1/m$ for all j, at the initialization of an optimization algorithm. The criterion $\phi_q(\xi_m;\mu_N)$ is then

$$\phi_q(\xi_m;\mu_N) = \left(\frac{1}{N} \sum_{i=1}^{N} \left[\sum_{j=1}^{m} w_j I_q(R_j,\rho_i) \right]^{-1} \right)^{1/q} \tag{19}$$

and the directional derivative of $\phi_q^q(\cdot;\mu_N)$ at ξ_m in the direction δ_{R_j} is $F_q(\xi_m,R_j;\mu_N) = \phi_q(\xi_m;\mu_N) - d(\xi_m,R_j;\mu_N)$, see (8), with

$$d(\xi_m,R_j;\mu_N) = \frac{1}{N} \sum_{i=1}^{N} \left\{ \left[\sum_{\ell=1}^{m} w_\ell I_q(R_\ell,\rho_i) \right]^{-2} I_q(R_j,\rho_i) \right\} .$$

The algorithms of Section 4.2 can then be applied straightforwardly.

4.4 Numerical examples

Example 3
Consider $\mathscr{X} = [-1,1]$ ($d = 1$) with $q = 1/4$. We take μ_N uniform on the first N points of Sobol' sequence (renormalized to $\mathscr{X}$) and $\xi_m^{(0)}$ is uniform on $s_{4N+1},\ldots,s_{4N+m}$ in the same sequence, with $N = m = 1\,024$. Figure 2-left presents the approximate density of the optimal measure obtained with the multiplicative algorithm (18) with $\lambda = 1$, stopped when $\max_i d(\xi_m^{(k)},z_i;\mu_N) - \phi_q^q(\xi_m^{(k)};\mu_N) < 10^{-4}$, see Theorem 2. Figure 2-right shows $d(\xi_m^{(k)},z;\mu_N)$ as a function of $z \in \mathscr{X}$ when the algorithm is stopped; the horizontal dashed line indicates the value of $\phi_q^q(\xi_m^{(k)};\mu_N)$.

Example 4
We take $\mathscr{X} = [0,1]^2$ ($d = 2$) and $q = 1/2$, and construct an approximation of the optimal design measure $\xi^{q,*}$, using μ_N given by the uniform distribution on the first N points $\mathbf{s}_1,\ldots,\mathbf{s}_N$ of Sobol' sequence in $\mathscr{X}$ and considering discrete measures ξ_m supported on $\mathbf{z}_1,\ldots,\mathbf{z}_m = \mathbf{s}_{4N+1},\ldots,\mathbf{s}_{4N+m}$ in the same sequence, with $N = m = 4\,096$. We perform $k = 300$ iterations of the multiplicative algorithm (18) with $\lambda = 1$ initialized with $\xi_m^{(0)}$ equal to the uniform measure on the $\mathbf{z}_i$. Figure 3-left presents the level sets corresponding to the 0.01, 0.025, 0.05 and 0.1 percentiles of the (smoothed) design mesure $\xi_m^{(k)}$. Figure 3-right shows two level sets $d(\xi_m^{(k)},\mathbf{z};\mu_N) = \phi_q^q(\xi_m^{(k)};\mu_N) - \varepsilon$ for $\varepsilon = 0.002$ and 0.001, indicating that almost all the mass of $\xi_m^{(k)}$ is allocated to points $\mathbf{z}_i$ such that $d(\xi_m^{(k)},\mathbf{z}_i;\mu_N)$ is very close to $\phi_q^q(\xi_m^{(k)};\mu_N)$ and that $\xi_m^{(k)}$ is close to being optimal (in fact, $\max_i d(\xi_m^{(k)},\mathbf{z}_i;\mu_N) - \phi_q^q(\xi_m^{(k)};\mu_N) < 8.8\;10^{-4}$, see Theorem 2).

Example 5
Take $\mathscr{X} = \mathscr{B}_d(0,1)$. Figure 4 shows $\phi_q^q(\delta_0)$, with δ_0 the delta measure at 0 (dotted line), $\phi_q^q(\mu)$

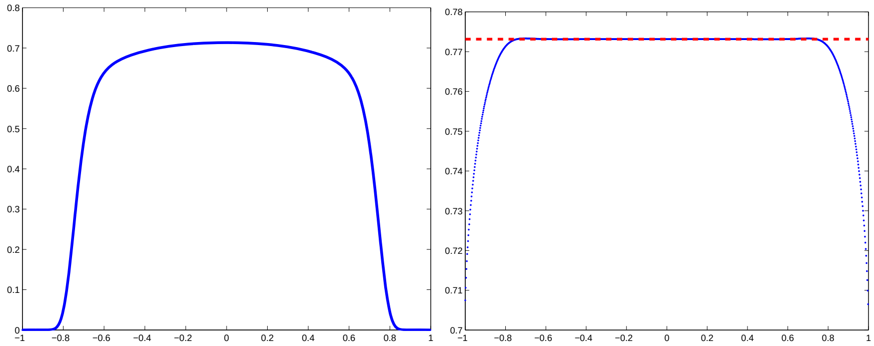

Fig. 2 Left: optimal design mesure $\xi_m^{q,*}$ (smoothed density). Right: $d(\xi_m^{q,*}, z; \mu_N)$ as a function of z and value of $\phi_q^q(\xi_m^{q,*}; \mu_N)$ (dashed line). $\mathscr{X} = [-1,1]$, $q = 1/4$, μ_N and ξ_m are supported on $m = N = 1\,024$ points of Sobol' sequence in $\mathscr{X}$.

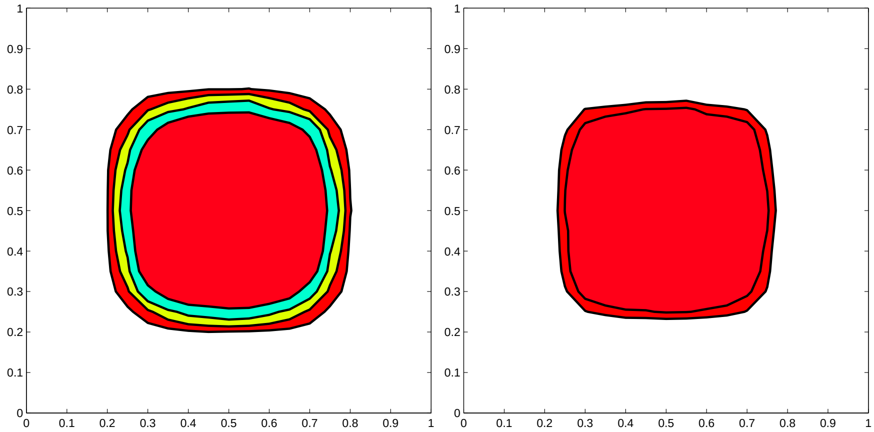

Fig. 3 Left: level sets corresponding to the 0.01, 0.025, 0.05 and 0.1 percentiles of the (smoothed) design mesure $\xi_m^{(k)}$. Right: levels sets $d(\xi_m^{(k)}, \mathbf{z}; \mu_N) = \phi_q^q(\xi_m^{(k)}; \mu_N) - 0.002$ and $d(\xi_m^{(k)}, \mathbf{z}; \mu_N) = \phi_q^q(\xi_m^{(k)}; \mu_N) - 0.001$. $\mathscr{X} = [0,1]^2$, $q = 1/2$, μ_N and ξ_m are supported by $m = N = 4096$ points of Sobol' sequence in $\mathscr{X}$, $\xi_m^{(k)}$ is obtained by $k = 300$ iterations (18) initialized with the uniform measure.

(dashed line), and an approximation of the optimal value $(\hat{\phi}_q^*)^q = \min\{\phi_q^q(\delta_0), \phi_q^q(\mu), \phi_q^q(\xi_m^{q,*}; \mu_N)\}$ (solid line), as functions of q in $[0,d]$, for $d = 3$ (left) and $d = 5$ (right). Here, $\xi_m^{q,*}$ minimizes $\phi_q(\xi_m; \mu_N)$, see (19), and is constructed with the vertex-exchange algorithm of Section 4.2.2. The radii ρ_i and R_j are given by 1 024 points of Sobol' sequence and the algorithm is stopped when $\max_j d(\xi_m^{(k)}, R_j; \mu_N) - \phi_q^q(\xi_m^{(k)}; \mu_N) < 10^{-4}$. As shown in Example 1, δ_0 is optimal for $0 < q < d-2$, with $\phi_q^q(\delta_0) = d/(d+q)$.

Figure 4 indicates that the optimal value ϕ_q^* is only marginally better than the value $\min\{\phi_q(\delta_0), \phi_q(\mu)\}$, hence the interest of considering the performance of the measure $\mu^{(r)}$ uniform on the ball $\mathscr{B}_d(0,r)$, with $\mu^{(0)} = \delta_0$ and $\mu^{(1)} = \mu$. In particular, note that the uniform measure μ is almost optimal for $q \in (d-1,d)$.

Using the expressions in Appendix, $\phi_q(\mu^{(r)})$ can be evaluated by numerical integration for $d = 3$ and $d = 5$ for any r and q, some analytical expressions being available for particular values of q and d, such as $\phi_1(\mu^{(r)}) = 3/4 - (27/4)\,r^4 + 6\,r^4\sqrt{3}\,\mathrm{arctanh}\left(\sqrt{3}/3\right)$ for $d = 3$.

Figure 5-left presents the efficiency $(\hat{\phi}_q^*)^q/\phi_q^q(\mu^{(r)})$ as a function of $r \in [0,1]$ when $d = 3$, for $q = 1/2$ (solid line) and $q = 3/2$ (dashed line). Figure 5-right shows the value of the radius r_q^* (obtained by numerical optimization) minimizing $\phi_q(\mu^{(r)})$ for $d = 3$ and $q \in [0,3]$. Numerical

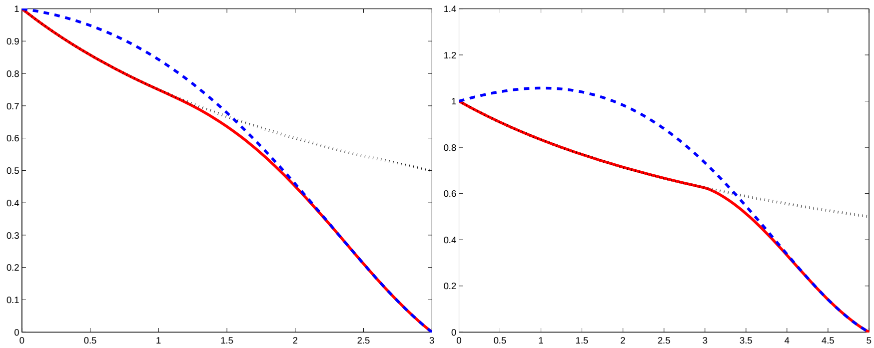

Fig. 4 $\phi_q^q(\delta_0)$ (dotted line), $\phi_q^q(\mu)$ (dashed line), and (approximate) optimal value $(\hat{\phi}_q^*)^q$ (solid line) as functions of q in $[0,d]$; $d=3$ (left) and $d=5$ (right).

calculations give $(\hat{\phi}_q^*)^q/\phi_q^q(\mu^{(r_q^*)}) > 99.9\%$ for all q in $[0,3]$, with $r_q^*=0$ for $q \leq 1$ and $r_q^* \simeq 1$ for $q \geq 5/2$, indicating that the uniform measure $\mu^{(r_q^*)}$ on $\mathscr{B}_3(0,r_q^*)$ is almost optimal for all q.

On the other hand, the shape of the optimal measure $\xi^{q,*}$ differs significantly from $\mu^{(r_q^*)}$ when $q \in (1,3)$, as shown on Figure 6 which presents the optimal density $\omega_{\xi^{q,*}}(\mathrm{d}r)/(dr^{d-1})$ for four different values of q (note that $\omega_{\mu^{(r_q^*)}}(\mathrm{d}r)/[dr^{d-1}/(r_q^*)^d]=1$ for $r \in [0,r_q^*]$ and equals zero elsewhere). Here $\xi^{q,*}$ is approximated using the technique in Section 4.3 and the multiplicative algorithm (18), with 4096 points from Sobol' sequence for the R_j and ρ_i in (19). The algorithm is stopped when $\max_j d(\xi_m^{(k)},R_j;\mu_N) - \phi_q^q(\xi_m^{(k)};\mu_N) < 10^{-4}$.

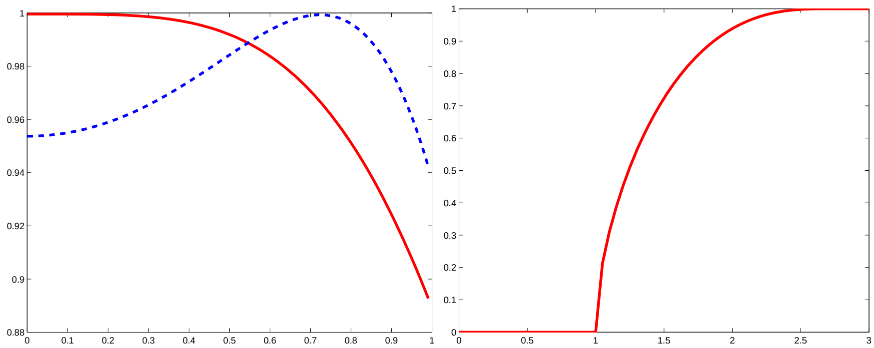

Fig. 5 Left: efficiency $(\hat{\phi}_q^*)^q/\phi_q^q(\mu^{(r)})$ as a function of $r \in [0,1]$ when $d=3$, for $q=1/2$ (solid line) and $q=3/2$ (dashed line). Right: r_q^* minimizing $\phi_q(\mu^{(r)})$ for $d=3$ as a function of $q \in [0,3]$.

5 Construction of n-point designs

As shown in Section 3.2, the optimal measure $\xi^{q,*}$ has no atoms when $\max\{0,d-2\} < q < d$. Also, when $q \geq d$, $\phi_q(\xi)=0$ for any measure equivalent to the Lebesgue measure, whereas $\phi_q(\xi) > 0$ when ξ is discrete, see Section 2.1. In this section we consider a method for generating n-point sets with small dispersion based on an algorithm for the construction of optimal design measures for $q>0$, and compare its performance with that of a greedy algorithm for the minimization of the coverage criterion (2).

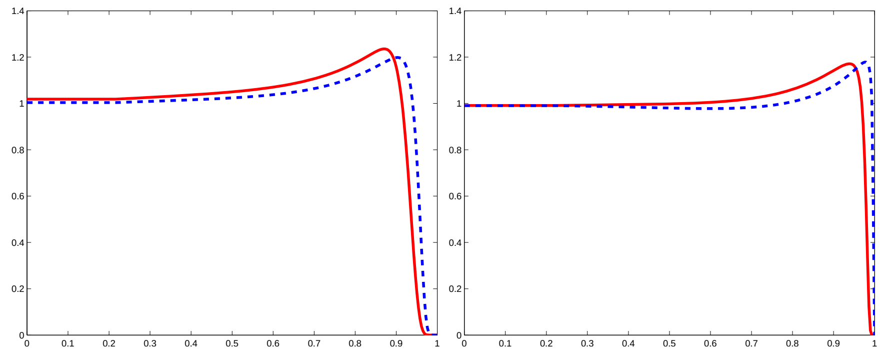

Fig. 6 Optimal density $\omega_{\xi_{q,*}}(\mathrm{d}r)/(dr^{d-1})$ for $d=3$: left for $q=2$ (solid line) and $q=2.1$ (dashed line); right for $q=2.25$ (solid line) and $q=2.5$ (dashed line).

5.1 A vertex-direction algorithm with predefined step-sizes

As in Section 4.2, we use a discrete approximation $\mu_N = (1/N)\sum_{i=1}^N \delta_{\mathbf{x}_i}$ of μ and consider the construction of design measures supported on a m-point set $\mathscr{Z}_m = \{\mathbf{z}_1,\ldots,\mathbf{z}_m\} \subset \mathscr{X}$, disjoint from $\mathscr{X}_N = \{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$, with N and m large enough (we assume that $m \gg n$, with m and N being of the same order of magnitude). But in contrast with Section 4.2, here we use this construction to generate n-point designs, for a given n.

Let X_{k_0} denote an initial k_0-point set in $\mathscr{Z}_m$. Denote by $\xi^{(k_0)}$ the corresponding empirical measure, $\xi^{(k_0)} = (1/k_0)\sum_{\mathbf{s}\in X_{k_0}} \delta_{\mathbf{s}}$, and consider the following vertex-direction algorithm with predefined step-sizes, initially proposed by Wynn [21] for D-optimal design: for all $k \geq k_0$, set

$$\xi^{(k+1)} = \left(1 - \frac{1}{k+1}\right)\xi^{(k)} + \frac{1}{k+1}\delta_{\mathbf{s}_{k+1}}\,, \quad \text{with } \mathbf{s}_{k+1} = \arg\max_{\mathbf{z}_\ell\in\mathscr{Z}_m} d(\xi^{(k)},\mathbf{z}_\ell;\mu_N) \qquad (20)$$

the best vertex-direction.

Denote by $\nabla_q^2(\xi,\mathbf{y})$ the second-order derivative $\mathrm{d}^2\phi_q^q[(1-\alpha)\xi+\alpha\delta_{\mathbf{y}}]/\mathrm{d}\alpha^2\big|_{\alpha=0}$ and by $\nabla_q^2(\xi,\mathbf{y};\mu_N)$ the value obtained when μ_N is substituted for μ. We obtain

$$\nabla_q^2(\xi,\mathbf{y}) = 2\int_{\mathscr{X}} \left\{ \frac{1}{P_{\xi,q}(\mathbf{x})} \left[1 - \frac{\|\mathbf{y}-\mathbf{x}\|^{-q}}{P_{\xi,q}(\mathbf{x})}\right]^2 \right\} \mu(\mathrm{d}\mathbf{x})\,,$$

with $P_{\xi,q}(\mathbf{x})$ given by (6). Although $\nabla_q^2(\xi,\mathbf{y})$ is infinite for $d/2 \leq q < d$ (it is finite for μ-almost all $\mathbf{y}$ when $q < d/2$, see Section 2.2), $\nabla_q^2(\xi,\mathbf{z}_j;\mu_N)$ is finite for any $q>0$ and all $j=1,\ldots,m$ when $\mathscr{Z}_m$ and $\mathscr{X}_N$ are disjoint. By standard arguments (see for instance [18, Chap. 9]), the convexity of $\phi_q^q(\cdot;\mu_N)$ and the non-summability of the sequence $\{1/k\}$ imply that $\phi_q(\xi^{(k)};\mu_N)$ tends to $\max_{\xi\in\Xi_m}\phi_q(\xi;\mu_N)$ as $k\to\infty$, with Ξ_m denoting the set of probability measures on $\mathscr{Z}_m$. Note that this property only relies on convexity, and is thus also valid for $q \geq d$, see Section 2.1. Values of q larger that d will be considered in Section 5.3. The convergence of (20) to an optimal measure is usually much slower than that of the algorithms in Section 4.2, but (20) provides a simple method for the generation of n-point sets.

When $q > d-2$ and m and N are large enough compared to n, the measure $\xi^{(k)}$ generated by (20) has more than n distinct support points for large k. We can thus construct an n-point set X_n by stopping (20) at iteration $k = k_n \geq n$, when $\xi^{(k_n)}$ has exactly n distinct support points (which often gives $k_n = n$). When the initial design X_{k_0} is suitably chosen (for instance, $k_0 = 1$ and $X_1 = \{\mathbf{s}_1\}$ with $\mathbf{s}_1 = \arg\min_{\mathbf{z}\in\mathscr{Z}_m}\sum_{i=1}^N \|\mathbf{z}-\mathbf{x}_i\|^q$, see (4)), in general the design X_n has already good space-filling properties, see Section 5.3. It can also be improved by local minimization, using for instance a Fedorov-type exchange algorithm [6, Chap. 3] for the minimization of $\Phi_q(\cdot;\mu_N)$ given by (2), similarly to [20].

One may notice that the calculation of the $d(\xi^{(k)},\mathbf{z}_\ell;\mu_N)$ in (20) only involves the evaluations of weighted sums of distances $\|\mathbf{x}_i-\mathbf{z}_\ell\|^{-q}$, since

$$d(\xi^{(k)},\mathbf{z}_\ell;\mu_N)=\frac{1}{N}\sum_{i=1}^N \nu_k(\mathbf{x}_i)\,\|\mathbf{x}_i-\mathbf{z}_\ell\|^{-q}\,,$$

with

$$\nu_k(\mathbf{x}_i)=\left(\sum_{j=1}^m \xi^{(k)}(\mathbf{z}_j)\,\|\mathbf{x}_i-\mathbf{z}_j\|^{-q}\right)^{-2}. \qquad (21)$$

All the terms $\|\mathbf{x}_i-\mathbf{z}_\ell\|^{-q}$, $i=1,\ldots,N$, $\ell=1,\ldots,m$, can be computed in advance, and we only need to update the sum in (21) to be able to determine $\mathbf{s}_{k+1}$ in (20).

Remark 2 *Let ξ be any measure in the set Ξ_m of probability measures on $\mathscr{Z}_m$, denote by $\xi^{q,*}$ an optimal measure that minimizes $\phi_q(\cdot;\mu_N)$ over Ξ_m and by $X_n^{q,*}$ an optimal n-point design minimizing $\Phi_q(\cdot;\mu_N)$ over $\mathscr{Z}_m$. We then have*

$$\begin{aligned}\Phi_q^q(X_n^{q,*};\mu_N)\geq\phi_q^q(\xi^{q,*};\mu_N)\geq\ &\phi_q^q(\xi;\mu_N)+\min_{\mathbf{z}_\ell\in\mathscr{Z}_m}F_q(\xi,\mathbf{z}_\ell;\mu_N)\\ =\ &2\phi_q^q(\xi;\mu_N)-\max_{\mathbf{z}_\ell\in\mathscr{Z}_m}d(\xi,\mathbf{z}_\ell;\mu_N)\,,\end{aligned}$$

see (10)*, where the first inequality follows from the optimality of $\xi^{q,*}$ and the second from the convexity of $\phi_q^q(\cdot;\mu_N)$. Therefore, any n-point design X_n in $\mathscr{Z}_m$ (in particular, an n-point design generated with* (20)*) satisfies*

$$\mathrm{eff}_q(X_n;\mu_N)=\frac{\Phi_q(X_n^{q,*};\mu_N)}{\Phi_q(X_n;\mu_N)}\geq B_q(X_n,\xi;\mu_N)\,,$$

where

$$B_q(X_n,\xi;\mu_N)=\left[\max\left\{0,2-\frac{\max_{\mathbf{z}_\ell\in\mathscr{Z}_m}d(\xi,\mathbf{z}_\ell;\mu_N)}{\phi_q^q(\xi;\mu_N)}\right\}\right]^{1/q}\frac{\phi_q(\xi;\mu_N)}{\Phi_q(X_n;\mu_N)}\,.$$

This lower bound on $\mathrm{eff}_q(X_n;\mu_N)$ is generally not informative when $\xi=\xi_n$, the empirical measure associated with X_n, in the sense that $B_q(X_n,\xi_n;\mu_N)$ is often equal to zero. On the other hand, using the algorithms of Section 4.2, we can easily obtain an ε-optimal measure $\xi_\varepsilon\in\Xi_m$ such that $\max_{\mathbf{z}_\ell\in\mathscr{Z}_m}d(\xi_\varepsilon,\mathbf{z}_\ell;\mu_N)/\phi_q^q(\xi_\varepsilon;\mu_N)<1+\varepsilon$, for a small ε, which gives

$$\mathrm{eff}_q(X_n;\mu_N)\geq B_q(X_n,\xi_\varepsilon;\mu_N)>(1-\varepsilon)^{1/q}\,\frac{\phi_q(\xi_\varepsilon;\mu_N)}{\Phi_q(X_n;\mu_N)}\,. \qquad (22)$$

5.2 Greedy minimization of the coverage criterion

The greedy minimization of the coverage criterion $\Phi_q(\cdot;\mu_N)$ given by (2) (with $q>0$) corresponds to

$$X_{k+1}=X_k\cup\hat{\mathbf{z}}_{k+1}\ \text{with}\ \hat{\mathbf{z}}_{k+1}=\arg\min_{\mathbf{z}_j\in\mathscr{Z}_m}\Phi_q(X_k\cup\{\mathbf{z}_j\};\mu_N)\,, \qquad (23)$$

for all $k\geq k_0$, for some given k_0-point set X_{k_0}. Note that this construction may involve repetitions; that is, some points can be repeated several times in X_k (including X_{k_0}). When it happens, like in Section 5.1 we continue the iterations until the design contains n distinct points. Also note that the determination of $\hat{\mathbf{z}}_{k+1}$ requires the exponentiation of $N\times m$ elements, whereas only N exponentiations are required for the calculation of $\mathbf{s}_{k+1}$ in (20), see (21).

Remark 3 *One can easily check that for any $q > 0$, the un-normalized version of $\Phi_q^q(\cdot;\mu_N)$ defined by $\tilde{\Phi}_q^q(X_n) = (1/n)\Phi_q^q(X_n;\mu_N)$ is non-increasing (that is, $\tilde{\Phi}_q^q(X_n \cup \{\mathbf{z}_j\}) \leq \tilde{\Phi}_q^q(X_n)$ for any X_n and $\mathbf{z}_j$ in $\mathscr{Z}_m$), and super-modular (that is, $\tilde{\Phi}_q^q(X_n) - \tilde{\Phi}_q^q(X_n \cup \{\mathbf{z}_j\}) \geq \tilde{\Phi}_q^q(X_n \cup \{\mathbf{z}_\ell\}) - \tilde{\Phi}_q^q(X_n \cup \{\mathbf{z}_\ell, \mathbf{z}_j\})$ for any X_n and $\mathbf{z}_j, \mathbf{z}_\ell$ in $\mathscr{Z}_m$). The selection of $\hat{\mathbf{z}}_{k+1}$ in* (23) *can also be written as*

$$\hat{\mathbf{z}}_{k+1} = \arg\min_{\mathbf{z}_j \in \mathscr{Z}_m} \tilde{\Phi}_q^q(X_k \cup \{\mathbf{z}_j\}) \,.$$

The results in [13] on the maximization of submodular set functions thus imply that the greedy minimization (23) *ensures that*

$$\frac{\frac{1}{k_0}\,\Phi_q^q(X_{k_0};\mu_N) - \frac{1}{k_0+t}\,\Phi_q^q(X_{k_0} \cup \{\hat{\mathbf{z}}_{k_0+1},\ldots,\hat{\mathbf{z}}_{k_0+t}\};\mu_N)}{\frac{1}{k_0}\,\Phi_q^q(X_{k_0};\mu_N) - \frac{1}{k_0+t}\,\Phi_q^q(X_{k_0} \cup \{\mathbf{z}^*_{k_0+1},\ldots,\mathbf{z}^*_{k_0+t}\};\mu_N)} \geq 1 - \left(1 - \frac{1}{t}\right)^t \geq 1 - 1/\mathrm{e}\,, \qquad (24)$$

for any non-empty set X_{k_0} and any $t \geq 1$, where the t points $\mathbf{z}^*_{k_0+1},\ldots,\mathbf{z}^*_{k_0+t}$ minimize $\Phi_q(X_{k_0} \cup \{\mathbf{z}_{k_0+1},\ldots,\mathbf{z}_{k_0+t}\};\mu_N)$ with respect to $\mathbf{z}_{k_0+1},\ldots,\mathbf{z}_{k_0+t}$ in $\mathscr{Z}_m$. Such efficiency bounds generally motivate the use of greedy algorithms for the construction of n-point sets with suitable properties; one should notice, however, that (24) *does not provide any exploitable bound on $\Phi_q^q(X_{k_0} \cup \{\mathbf{z}^*_{k_0+1},\ldots,\mathbf{z}^*_{k_0+t}\};\mu_N)/\Phi_q^q(X_{k_0} \cup \{\hat{\mathbf{z}}_{k_0+1},\ldots,\hat{\mathbf{z}}_{k_0+t}\};\mu_N)$.*

5.3 Numerical examples

Example 6
We take $\mathscr{X} = [0,1]^2$, with $\mathscr{Z}_m$ formed by the 33×33 regular grid with $\{\mathbf{z}_j\}_{1,2} \in \{0, 1/32, 2/32, \ldots, 1\}$, and $\mathscr{X}_N$ formed by the 32×32 regular grid interlaced with $\mathscr{Z}_m$, with $\{\mathbf{x}_i\}_{1,2} \in \{1/64, 1/64 + 1/32, \ldots, 1/64 + 31/32\}$.

Figure 7 shows the 100-point designs generated by (20) with $X_1 = \{(1/2, 1/2)\}$ for $q = 1$ (left) and $q = 3/2$ (right), illustrating the better space-filling behaviour of the design generated when q increases.

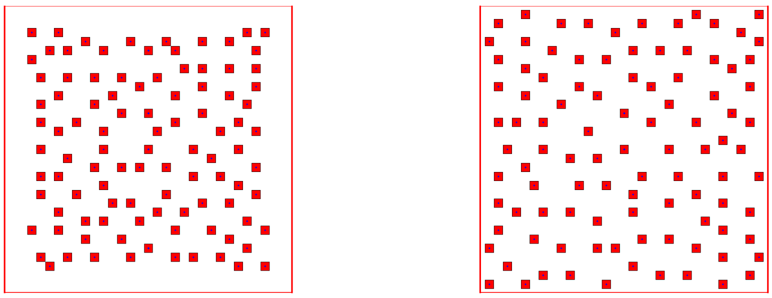

Fig. 7 Designs X_{100} generated by (20) for $q = 1$ (left) and $q = 3/2$ (right).

Figure 8-left shows the design X_{50}^a generated by (20) with $X_1 = \{(1/2, 1/2)\}$ for $q = 10$. The lower bound (22) with $\varepsilon = 10^{-4}$ gives $\mathrm{eff}_q(X_{50}^a;\mu_N) > 0.3292$. We also obtain $\mathrm{eff}_q(\mathscr{Z}_m;\mu_N) > 0.9938$, which indicates that the uniform measure on $\mathscr{Z}_m$ is almost optimal. The dispersion of X_{50}^a (evaluated by Voronoï tessellation, see [16, Section 2.4]) is $\Phi(X_{50}^a) \simeq 0.1141$; its maximin-distance is $\Psi(X_n^{(a)}) \simeq 0.0938$, see (17). The design X_{50}^b in Figure 8-right is generated by (23), with the same X_1 and the same q, and $\Phi(X_{50}^b) \simeq 0.1310$, $\Psi(X_n^{(b)}) \simeq 0.0699$. The designs obtained can be used to initialize a local optimization algorithm for the minimization of $\Phi_q(\cdot;\mu_N)$.

In particular, the small values of m and n used in the example allow us to apply the Fedorov-type exchange algorithm proposed in [20]. When initialized at X_{50}^a (respectively X_{50}^b), it yields a design $\tilde{X}_{50}^a$ with $\Phi(\tilde{X}_{50}^a) \simeq 0.1127$ and $\Psi(\tilde{X}_n^{(a)}) \simeq 0.0988$ (respectively, $\tilde{X}_{50}^b$ with $\Phi(\tilde{X}_{50}^b) \simeq 0.1127$ and $\Psi(\tilde{X}_n^{(b)}) \simeq 0.0938$).

Performances are slightly worse for smaller values of q, with for instance $\Phi(X_{50}^a) \simeq 0.1273$ and $\Phi(X_{50}^b) \simeq 0.1563$ when $q = 2$.

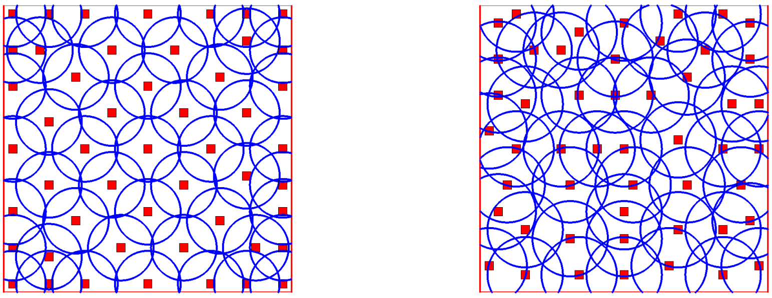

Fig. 8 Left: design X_{50}^a generated by (20); Right: X_{50}^b generated by (23); $X_1 = \{(1/2,1/2)\}$, $q = 10$. The circles have radii equal to the dispersion, with $\Phi(X_{50}^a) \simeq 0.1141$ and $\Phi(X_{50}^b) \simeq 0.1310$.

Algorithm (20) does not require $\mathscr{X}$ to be convex. Figure 9-left shows the design obtained when $\mathscr{X} = [0,1]^2 \setminus \mathscr{B}_2((1/2,1/2),1/4)$ and $q = 10$. The set $\mathscr{X}_m$ (respectively, $\mathscr{X}_N$) corresponds to the subset of the first m' points $\mathbf{s}_1,\ldots,\mathbf{s}_{m'}$ of Sobol' sequence in $[0,1]^2$ (respectively, the subset of $\mathbf{s}_{4m'+1},\ldots,\mathbf{s}_{4m'+N'}$) that lie in $\mathscr{X}$, with $m' = N' = 4096$ (which gives $m = 3292$ and $N = 3295$). The direct use of the first 50 points of Sobol' sequence that fall in $\mathscr{X}$ gives the design in Figure 9-right, with clearly worse space-filling properties.

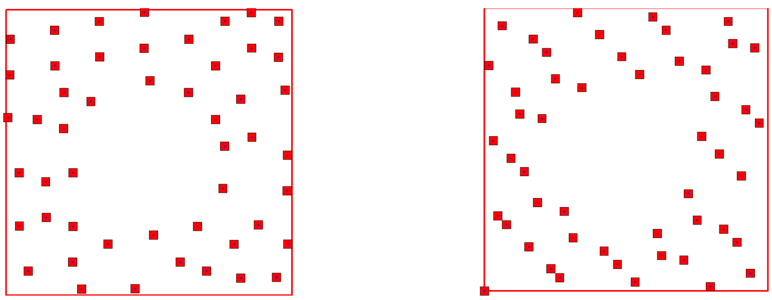

Fig. 9 Left: design X_{50} generated by (20) when $\mathscr{X} = [0,1]^2 \setminus \mathscr{B}_2((1/2,1/2),1/4)$; Right: first 50 points of Sobol' sequence in $\mathscr{X}$.

Example 7

We take $\mathscr{X} = [0,1]^{10}$ and $n = 100$; $\mathscr{X}_m$ corresponds to the first m points $\mathbf{s}_1,\ldots,\mathbf{s}_m$ of Sobol' sequence in $\mathscr{X}$, $\mathscr{X}_N$ corresponds to $\mathbf{s}_{4m+1},\ldots,\mathbf{s}_{4m+N}$, with $m = N = 4096$. The design X_{100}^a generated by (20) with $q = 10$ has dispersion $\Phi(X_{100}^a) \simeq 1.310$ (computed with the MCMC method in [16, Section 2.4]) and $\Psi(X_{100}^a) \simeq 0.6011$. The lower bound (22) (with $\varepsilon = 10^{-4}$) gives $\mathrm{eff}_q(X_{100}^a; \mu_N) > 0.8288$ and $\mathrm{eff}_q(\mathscr{Z}_m; \mu_N) > 0.9702$. The computational time of X_n^a, including the computation of all $\|\mathbf{x}_i - \mathbf{z}_\ell\|^{-q}$, is about 2.8 s in Matlab on a PC with a clock speed of 2.50 Gz and 32 Go RAM. Algorithm (23) requires about 30 s to generate a design X_n^b with $\Phi(X_{100}^b) \simeq 1.349$ and $\Phi(X_{100}^b) \simeq 0.3550$. Similar behaviours are observed for the hypercube $[0,1]^d$ with different d and other values of m, N and n, and indicate that the vertex-direction algorithm (20) is significantly faster than the greedy-method (23) and yields designs with better space-filling properties.

6 Conclusions

We have shown (Theorem 1) that a continuous extension of a regularized version of the dispersion criterion is a convex functional when the regularization parameter q is positive, and strictly convex when $0 < q < d$. For the range of q where the functional is strictly convex, we have given a characterization of optimal measures (Theorem 2). Their properties have been investigated using results from potential theory and several examples have been studied in details. In particular, analytic results have been obtained for the case when the design space is the unit ball and experimental design algorithms have been used for numerical constructions. In the same way as the determination of Fekete points that minimize the q-energy is easier than the construction of an optimal design for the maximin criterion (17), see [2], [12], the minimization of (2) is easier that the construction of a design minimizing the dispersion (1). The results in the paper suggest that the construction of designs having good space-filling properties requires q to be larger than $d-2$, but that very large values of q are not needed. Numerical experiments indicate that n-point designs with good space-filling properties can easily be generated with a vertex-direction algorithm with predefined step-sizes.

Appendix: explicit expressions of integrals for $\mathscr{X} = \mathscr{B}_d(0,1)$

In this appendix we give the explicit expression of the integral (15), and also of (14) when $\xi = \mu$, for particular values of d and q when $\mathscr{X} = \mathscr{B}_d(0,1)$.

1. $q = d-2$, any $d \geq 2$

We have $I_{d,q}(R,\rho) = (\max\{R,\rho\})^{-q}$, implying $J_{d,q}(\rho) = \rho^2(1-d/2) + d/2$ when $\mathbf{z}$ is uniformly distributed in $\mathscr{B}_d(0,1)$ (so that $\omega_\xi(\mathrm{d}R) = dR^{d-1}\,\mathrm{d}R$ in (14)). Hence, the integral (13) can be evaluated exactly when $\xi = \mu$. Table 1 gives the values of $\phi_q^q(\mu)$ for $d = 3, \ldots, 10$ and $q = d-2$.

d	$\phi_{d-2}^{d-2}(\mu)$	
3	$6\sqrt{3}\,\mathrm{arctanh}\left(\sqrt{3}/3\right) - 6$	$\simeq 0.843113970$
4	$4\log(2) - 2$	$\simeq 0.772588722$
5	$[2\cdot 5^2\sqrt{15}\,\mathrm{arctanh}\left(\sqrt{15}/5\right)]/27 - 20/3$	$\simeq 0.733016111$
6	$3^3\log(3)/2^3 - 3$	$\simeq 0.707816475$
7	$[2\cdot 7^3\sqrt{35}\,\mathrm{arctanh}\left(\sqrt{35}/7\right)]/5^4 - 2758/375$	$\simeq 0.690405156$
8	$2^9\log(2)/3^4 - 100/27$	$\simeq 0.677671067$
9	$[2\cdot 3^9\sqrt{7}\,\mathrm{arctanh}\left(\sqrt{7}/3\right)]/7^5 - 94968/12005$	$\simeq 0.667959803$
10	$5^5\log(5)/2^{10} - 3265/768$	$\simeq 0.660312638$

Table 1 Values of $\phi_{d-2}^{d-2}(\mu)$ for $d = 3, \ldots, 10$ when $\mathscr{X} = \mathscr{B}_d(0,1)$.

2. $d=3$ and $0<q<3$

We obtain

$$I_{3,q}(R,\rho) = \frac{\left((R-\rho)^2\right)^{1-q/2} - \left((\rho+R)^2\right)^{1-q/2}}{2\rho\,(q-2)R} \quad \text{for } q \neq 2\,,$$
$$I_{3,2}(R,\rho) = \frac{\log\left((\rho+R)^2\right) - \log\left((R-\rho)^2\right)}{4\rho R} \quad \text{for } q=2\,.$$

When $\mathbf{z}$ is uniformly distributed in $\mathscr{B}_3(0,1)$, the evaluation of

$$J_{d,q}(\rho) = 3\int_0^1 I_{d,q}(R,\rho)R^2\,\mathrm{d}R$$

gives

$$J_{3,2}(\rho) = 3\,\frac{(1-\rho^2)\log[(1+\rho)/(1-\rho)]+2\rho}{4\rho}\,,$$
$$J_{3,q}(\rho) = 3\,\frac{(\rho+1)^{3-q}(q+\rho-3)-(q-\rho-3)(1-\rho)^{3-q}}{2\rho\,(q-2)(q-3)(q-4)} \quad \text{for } q\neq 2\,.$$

The numerical evaluation of the integral $\phi_q(\mu)$ is now straightforward, see (13); $\phi_1(\mu)$ can be computed exactly, see Table 1; also, $\lim_{q\to 0}\phi_q(\mu) = 2\exp(-3/4) \simeq 0.9447331$.

3. $d=5$ and $0<q<5$

If $q \notin \{2,3,4\}$ then

$$I_{5,q}(R,\rho) = 3\,\frac{(R+\rho)^\beta - \beta R\rho\,(R+\rho)^\alpha - |R-\rho|^\beta - \beta R\rho\,|R-\rho|^\alpha}{2R^3\rho^3\alpha\,(\alpha^2-4)}\,,$$

where $\alpha = 4-q > -1$ and $\beta = 6-q = \alpha+2$. The integral (14) explicitly evaluates when $\mathbf{z}$ is uniform in $\mathscr{B}_5(0,1)$, using the representation

$$J_{5,q}(\rho) = 5\int_0^1 I_{5,q}(R,\rho)R^4\,\mathrm{d}R = c\,(J_1 - \beta\rho J_2 - J_3 - \beta\rho J_4)\,,$$

where $c = 15/\left[2\rho^3\alpha\left(\alpha^2-4\right)\right]$ and

$$J_1 = \int_0^1 (R+\rho)^\beta R\,\mathrm{d}R = \frac{\rho^{\beta+2}+(\rho+1)^{\beta+1}(\beta-\rho+1)}{(\beta+2)(\beta+1)}\,,$$
$$J_2 = \int_0^1 (R+\rho)^\alpha R^2\,\mathrm{d}R = \frac{\left(\alpha^2-2\alpha\rho+2\rho^2+3\alpha-2\rho+2\right)(1+\rho)^{\alpha+1}-2\rho^{\alpha+3}}{(\alpha+3)(\alpha+2)(\alpha+1)}\,,$$
$$J_3 = \int_0^1 |R-\rho|^\beta R\,\mathrm{d}R = \frac{\rho^{\beta+2}+(\rho+\beta+1)(1-\rho)^{\beta+1}}{(\beta+1)(\beta+2)}\,,$$
$$J_4 = \int_0^1 |R-\rho|^\alpha R^2\,\mathrm{d}R = \frac{\left(\alpha^2+2\alpha\rho+2\rho^2+3\alpha+2\rho+2\right)(1-\rho)^{\alpha+1}+2\rho^{\alpha+3}}{(\alpha+3)(\alpha+2)(\alpha+1)}\,,$$

so that $\phi_q(\mu)$ can easily be evaluated numerically.

When $q = 2$ and $q = 4$, we obtain

$$I_{5,2}(R,\rho) = 3\,\frac{2\rho R(R^2+\rho^2)+(R^2-\rho^2)^2\log|(R-\rho)/(R+\rho)|}{16R^3\rho^3}\,,$$

$$J_{5,2}(\rho) = 5\,\frac{2\rho(3-\rho^2)(3\rho^2+1)+(1-\rho^2)^3\log[(1-\rho)/(1+\rho)]}{96\rho^3}\,,$$

$$I_{5,4}(R,\rho) = 3\,\frac{-2\rho R+(R^2+\rho^2)\log|(R+\rho)/(R-\rho)|}{8R^3\rho^3}\,,$$

$$J_{5,4}(\rho) = 15\,\frac{2\rho(3\rho^2-1)+(1-\rho^2)(1+3\rho^2)\log[(1-\rho)/(1+\rho)]}{32\rho^3}\,.$$

The case $q = 3$ is considered in Table 1.

Acknowledgements We thank the two anonymous referees for their comments that helped us to improve the presentation and incited us to consider the construction of n-point designs (Section 5).

References

1. Armijo, L.: Minimization of functions having Lipschitz continuous first partial derivatives. Pac. J. Math. **16**, 1–3 (1966)
2. Audze, P., Eglais, V.: New approach for planning out of experiments. Problems of Dynamics and Strengths **35**, 104–107 (1977)
3. Böhning, D.: Numerical estimation of a probability measure. Journal of Statistical Planning and Inference **11**, 57–69 (1985)
4. Böhning, D.: A vertex-exchange-method in D-optimal design theory. Metrika **33**, 337–347 (1986)
5. Conway, J., Sloane, N.: Sphere Packings, Lattices and Groups. Springer, New York (1999). [3rd ed.]
6. Fedorov, V.: Theory of Optimal Experiments. Academic Press, New York (1972)
7. Hardin, D., Saff, E.: Discretizing manifolds via minimum energy points. Notices of the AMS **51**(10), 1186–1194 (2004)
8. Johnson, M., Moore, L., Ylvisaker, D.: Minimax and maximin distance designs. Journal of Statistical Planning and Inference **26**, 131–148 (1990)
9. Landkof, N.: Foundations of Modern Potential Theory. Springer, Berlin (1972)
10. Molchanov, I., Zuyev, S.: Variational calculus in the space of measures and optimal design. In: A. Atkinson, B. Bogacka, A. Zhigljavsky (eds.) Optimum Design 2000, chap. 8, pp. 79–90. Kluwer, Dordrecht (2001)
11. Molchanov, I., Zuyev, S.: Steepest descent algorithm in a space of measures. Statistics and Computing **12**, 115–123 (2002)
12. Morris, M., Mitchell, T.: Exploratory designs for computational experiments. Journal of Statistical Planning and Inference **43**, 381–402 (1995)
13. Nemhauser, G., Wolsey, L., Fisher, M.: An analysis of approximations for maximizing submodular set functions–I. Mathematical Programming **14**(1), 265–294 (1978)
14. Niederreiter, H.: Random Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia (1992)
15. Pronzato, L.: A delimitation of the support of optimal designs for Kiefer's ϕ_p-class of criteria. Statistics & Probability Letters **83**, 2721–2728 (2013)
16. Pronzato, L.: Minimax and maximin space-filling designs: some properties and methods for construction. Journal de la Société Française de Statistique **158**(1), 7–36 (2017)
17. Pronzato, L., Müller, W.: Design of computer experiments: space filling and beyond. Statistics and Computing **22**, 681–701 (2012)
18. Pronzato, L., Pázman, A.: Design of Experiments in Nonlinear Models. Asymptotic Normality, Optimality Criteria and Small-Sample Properties. Springer, LNS 212, New York (2013)
19. Pronzato, L., Wynn, H., Zhigljavsky, A.: Extremal measures maximizing functionals based on simplicial volumes. Statistical Papers **57**(4), 1059–1075 (2016). Hal-01308116
20. Royle, J., Nychka, D.: An algorithm for the construction of spatial coverage designs with implementation in SPLUS. Computers and Geosciences **24**(5), 479–488 (1998)
21. Wynn, H.: The sequential generation of D-optimum experimental designs. Annals of Math. Stat. **41**, 1655–1664 (1970)
22. Yu, Y.: Monotonic convergence of a general algorithm for computing optimal designs. Annals of Statistics **38**(3), 1593–1606 (2010)