



**HAL**  
open science

## Incremental space-filling design based on coverings and spacings: improving upon low discrepancy sequences

Amaya Nogales Gómez, Luc Pronzato, Maria-João Rendas

► **To cite this version:**

Amaya Nogales Gómez, Luc Pronzato, Maria-João Rendas. Incremental space-filling design based on coverings and spacings: improving upon low discrepancy sequences. 2020. hal-02987983v1

**HAL Id: hal-02987983**

**<https://hal.science/hal-02987983v1>**

Preprint submitted on 4 Nov 2020 (v1), last revised 26 May 2021 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Incremental space-filling design based on coverings and spacings: improving upon low discrepancy sequences

A. Nogales Gómez, L. Pronzato, M.-J. Rendas

UCA, CNRS, Laboratoire I3S, UMR 7271,  
Bât. Euclide, Les Algorithmes, 2000 route des lucioles,  
06900 Sophia Antipolis cedex, France  
{amaya.nogales-gomez,pronzato,rendas}@i3s.unice.fr

November 4, 2020

## Abstract

The paper addresses the problem of defining families of ordered sequences  $\{\mathbf{x}_i\}_{i \in \mathbb{N}}$  of elements of a compact subset  $\mathcal{X}$  of  $\mathbb{R}^d$  whose prefixes  $\mathbf{X}_n \triangleq \{\mathbf{x}_i\}_{i=1}^n$ , for all orders  $n$ , have good space-filling properties as measured by the dispersion (covering radius) criterion. Our ultimate aim is the definition of incremental algorithms that generate sequences  $\mathbf{X}_n$  with small optimality gap, i.e., with a small increase in the maximum distance between points of  $\mathcal{X}$  and the elements of  $\mathbf{X}_n$  with respect to the optimal solution  $\mathbf{X}_n^*$ . The paper is a first step in this direction, presenting incremental design algorithms with proven optimality bound with respect to one-parameter families of criteria based on coverings and spacings that both converge to dispersion for large values of their parameter.

**keywords** Covering; Spacing; Submodularity; Greedy algorithm; Computer experiments; Space-filling design

**MSC** 62K99, 65D99

## 1 Introduction and motivation

The paper discusses algorithmic constructions to define space-filling designs. Given a compact subset  $\mathcal{X} \subset \mathbb{R}^d$ , we say that a finite subset  $\mathbf{Z}_n \subset \mathcal{X}$  is a space-filling design if  $\mathbf{Z}_n$  fills  $\mathcal{X}$  evenly. Several mathematical definitions of this intuitive notion have been proposed in the literature, and we refer the interested reader to the reviews [25, 26] and the books [4], [31, Chap. 5] for a comprehensive presentation and discussion. Generically, a space-filling criterion is a set function, that maps finite subsets of  $\mathcal{X}$  to  $\mathbb{R}$ . Herein we focus in the covering radius, a measure of the space-filling quality of  $\mathbf{Z}_n$  motivated by interpolation problems, equal to the maximum distance between a point of  $\mathcal{X}$  and the elements of  $\mathbf{Z}_n$ : good space-filling sequences having small covering radius. In colloquial terms, “these designs leave no large holes in  $\mathcal{X}$ ”.

The packing radius is in some sense a dual space-filling metric, assessing how much  $\mathbf{Z}_n$  is scattered through the distance between its closest points. In terms of this criterion, a design is space-filling if it has a large packing radius. It is easy to see that the designs optimal for the covering radius are not optimal according to the packing radius: the latter must have some

points in the boundary of  $\mathcal{X}$  (otherwise a larger packing radius would be obtained by expanding the design uniformly), and this clearly does not lead to minimal covering radius. Nevertheless, the packing radius is widely used as a space-filling criterion, given the much smaller numerical complexity involved in this computation when compared to the covering radius. Besides the central criterion of interest, the covering radius, we will also assess the packing quality of the designs produced by the algorithms presented in the paper, and consider their mesh-ratio, defined as the ratio of the covering and packing radii.

**Background.** Optimization of the covering radius is a highly non-linear and inherently multi-dimensional combinatorial problem: the cost function involves a maximum and the search space has dimension  $n \times d$ . Actually, the problem is known to be NP-hard, see [13, p. 414], meaning that except for toy-problems we can only hope to find tractable algorithms that produce reasonably good solutions. In these circumstances, it is utterly important to know how far the solutions found by a given algorithm can be from optimality, commonly designated in algorithmics by optimality gap. The definition of algorithms with guaranteed optimality gap for NP-hard problems is an active research topic which has produced important achievements for many combinatorial optimization problems, and our work is a contribution in this sense.

**Claims and hint of the contents.** Algorithms which define the solution by incrementally adding a point at a time are specially interesting given their small complexity: at each iteration an optimization problem in  $\mathcal{X}$ , i.e., in only  $d$  variables, needs to be solved. It is a well known — and easily verified — fact that besides additive cost functions, for which the optimal solution can be found by greedily appending the point in  $\mathcal{X}$  that produces the largest improvement, the optimality gap of greedy algorithms can be bounded if the criterion optimized is a submodular set function (this statement will be made more precise below). Unfortunately, neither the covering radius nor the packing radius are submodular set functions. In this paper, we present a parametrized design criterion which is asymptotically (for large values of its parameter) related to the covering radius and is submodular, its greedy solutions having thus a bounded optimality gap. We also consider the incremental construction of designs by greedy maximization of the packing radius, sometimes called coffee-house design in the literature [19, 20]. Interestingly enough, this simple algorithm ensures an optimality gap of 50% for the covering radius [8]. A slight modification of the original algorithm, based on the notion of spacings, permits to keep the design points away from the boundary of  $\mathcal{X}$ , thereby reducing the covering radius.

There exist other incremental constructions that generate space-filling designs with good covering performance, in particular those based on maximization of mutual information [2, 14] or on minimization of a kernel discrepancy by kernel herding [29]. They rely on the choice of a suitable positive definite kernel and are especially adapted to interpolation based on Gaussian process models. We do not follow this approach here, and the methods we consider are based only on geometrical considerations. A thorough comparison would certainly be of interest, but is beyond the scope of this paper.

**Paper organization.** This paper is based on results presented at the SIAM Conference on Uncertainty Quantification in 2016 [27]. It is organized as follows. In Section 2 we recall the definition of submodular set functions and the fundamental theorem that establishes a bound on the optimality gap of their greedy solutions. Section 3 presents a novel design criterion based

on the cover measure, addressing in particular issues concerning its numerical evaluation and illustrating its behaviour for several implementation choices. Other greedy constructions, based on packing radius and spacings, are presented in Section 4. Finally, Section 5 compares the performances of the methods presented, both in terms of the quality of the designs obtained by its greedy optimization and of their numerical complexity.

**Basic definitions and notation.** Throughout the paper,  $\mathcal{X}$  is a compact and convex subset of  $\mathbb{R}^d$  with nonempty interior, with the hypercube  $\mathcal{C}_d = [0, 1]^d$  as typical example.  $\mathbf{1}_d$  denotes the  $d$ -dimensional vector with all components equal to one, and the center of  $\mathcal{C}_d$  is thus  $\mathbf{1}_d/2$ . Let  $\mathbf{Z}_n = \{\mathbf{z}_1, \dots, \mathbf{z}_n\}$  be any  $n$ -point design in  $\mathcal{X}$ . For any point  $\mathbf{x} \in \mathcal{X}$ , we denote  $d(\mathbf{x}, \mathbf{Z}_n) = \min_{i=1, \dots, n} \|\mathbf{x} - \mathbf{z}_i\|$ , with  $\|\cdot\|$  the  $\ell_2$  norm. The covering radius  $\text{CR}(\mathbf{Z}_n)$  of  $\mathbf{Z}_n$  is defined by

$$\text{CR}(\mathbf{Z}_n) = \text{CR}_{\mathcal{X}}(\mathbf{Z}_n) = \max_{\mathbf{x} \in \mathcal{X}} d(\mathbf{x}, \mathbf{Z}_n).$$

It is called dispersion in the theory of quasi-Monte Carlo methods [22, Chap. 6] and corresponds to the minimax-distance, or fill criterion, used in computer experiments and function interpolation; see [3, 11, 25, 26]. When the design objective is to ensure a precise prediction of the values of an unknown function  $f$  over all  $\mathcal{X}$  based on evaluations at  $\mathbf{Z}_n \in \mathcal{X}$ , it is important to ensure that for any  $\mathbf{x}$  in  $\mathcal{X}$  there always exists a  $\mathbf{z}_i$  at proximity of  $\mathbf{x}$  where  $f(\mathbf{z}_i)$  has been evaluated (see, e.g., [32] for a precise formulation and error bounds), making designs with a small value of  $\text{CR}(\mathbf{Z}_n)$  particularly desirable.

We denote by  $\mathcal{X}_N = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N)}\}$  a finite subset of  $\mathcal{X}$ , with usually  $N \gg n$ , that is well spread over  $\mathcal{X}$ ; it may be a regular grid when  $\mathcal{X} = \mathcal{C}_d$  and  $d$  is small, or (possibly after a suitable rescaling of  $\mathcal{X}$ ) the first  $N$  points of a low discrepancy sequence in  $\mathcal{C}_d$  that fall in  $\mathcal{X}$ . With a slight abuse of notation, we assume that  $\mathcal{X}_N$  and  $\mathcal{X}_{N'}$  do not necessarily share any elements when  $N \neq N'$ . Note that the value

$$\text{CR}_{\mathcal{X}_N}(\mathbf{Z}_n) = \max_{\mathbf{x} \in \mathcal{X}_N} d(\mathbf{x}, \mathbf{Z}_n) \tag{1}$$

underestimates  $\text{CR}(\mathbf{Z}_n)$ , the amount underestimation depending how well  $\mathcal{X}_N$  is spread in  $\mathcal{X}$ . In particular, we have

$$\text{CR}_{\mathcal{X}_N}(\mathbf{Z}_n) \leq \text{CR}(\mathbf{Z}_n) \leq \max_{\mathbf{x} \in \mathcal{X}} \min_{\mathbf{z}_i \in \mathbf{Z}_n, \mathbf{x}^{(j)} \in \mathcal{X}_N} \left( \|\mathbf{x} - \mathbf{x}^{(j)}\| + \|\mathbf{x}^{(j)} - \mathbf{z}_i\| \right) \leq \text{CR}_{\mathcal{X}_N}(\mathbf{Z}_n) + \text{CR}(\mathcal{X}_N). \tag{2}$$

The packing radius, or separating radius, or maximin-distance,

$$\text{PR}(\mathbf{Z}_n) = \min_{\mathbf{z}_i \neq \mathbf{z}_j \in \mathbf{Z}_n} \frac{1}{2} \|\mathbf{z}_i - \mathbf{z}_j\| \quad (n \geq 2),$$

is also often used as space-filling characteristic for a design  $\mathbf{Z}_n$  [3, 11, 25, 26], in particular due to the simplicity of its calculation when compared to  $\text{CR}(\mathbf{Z}_n)$ . Good designs should have a large packing radius. Notice, however, that  $\text{PR}(\mathbf{Z}_n)$  is a more local space-filling characteristic of  $\mathbf{Z}_n$  than  $\text{CR}(\mathbf{Z}_n)$ , in the sense that moving only one point  $\mathbf{z}_i$  of  $\mathbf{Z}_n$  to make it coincide with another  $\mathbf{z}_j$  sets  $\text{PR}(\mathbf{Z}_n)$  to zero whereas  $\text{CR}(\mathbf{Z}_n)$  only increases to  $\text{CR}(\mathbf{Z}_n \setminus \{\mathbf{z}_i\})$ .

We reserve the notation  $\mathbf{X}_n = [\mathbf{x}_1, \dots, \mathbf{x}_n]$  to designs that are constructed incrementally, our objective being that all prefix designs  $\mathbf{X}_m$ ,  $m = 1, \dots, n$ , should have good space-filling properties. We are especially interested in situations where  $n$  is small, that is, much smaller

than  $2^d$  when  $d$  is large, having in particular in mind the “10  $d$ ” rule of [16]. Restriction to incremental constructions adds an important complexity, since designs with minimum covering radius are not nested. The consequences are well known in the large  $n$  situation: for  $d = 1$ , remember for instance the low dispersion sequence of Ruzsa whose discrepancy does not tend to zero as  $n \rightarrow \infty$  [22, p. 154]. But the phenomenon is already present for small  $n$ : for  $\mathcal{X} = \mathcal{C}_d$  the one-point CR-optimal design is  $\mathbf{Z}_1^* = \{\mathbf{1}_d/2\}$  (the center of  $\mathcal{C}_d$ ), whereas 2-point CR-optimal designs have the form  $\mathbf{Z}_2^* = \{\mathbf{z}_1^*, \mathbf{z}_2^* = \mathbf{1}_d - \mathbf{z}_1^*\}$  with  $\mathbf{z}_1^*$  having all its coordinates equal to  $1/2$  except one which equals  $1/4$ ; see the Appendix. Any incremental construction is therefore already suboptimal for  $\text{CR}(\mathbf{X}_n)$  at  $n = 2$  (and  $\text{CR}(\mathbf{X}_n)$  can remain equal to  $\sqrt{d}/2$  for some iterations when the first design points are chosen near the vertices of  $\mathcal{C}_d$ ; see [5]).

## 2 Performance guarantee for submodular function maximization

This section recalls the notion of submodular set functions and a well known result establishing, when they are monotone, bounds on the optimality gap of their greedy maximization.

### 2.1 Submodularity and the greedy algorithm

Let  $\mathcal{X}_C$  be a finite candidate set with  $C = |\mathcal{X}_C|$  elements and denote by  $2^{\mathcal{X}_C}$  its power set. A (scalar real valued) set-function is an application of  $2^{\mathcal{X}_C}$  on  $\mathbb{R}$ . Space-filling criteria are set-functions.

**Definition 1 (Submodular function)** *A set-function  $f : 2^{\mathcal{X}_C} \rightarrow \mathbb{R}$  is submodular if and only if it satisfies the following three equivalent conditions, see e.g., [1].*

- (a)  $f(\mathbf{A}) + f(\mathbf{B}) \geq f(\mathbf{A} \cup \mathbf{B}) + f(\mathbf{A} \cap \mathbf{B}), \quad \forall \mathbf{A}, \mathbf{B} \in 2^{\mathcal{X}_C};$
- (b)  $f(\mathbf{A} \cup \{\mathbf{x}\}) - f(\mathbf{A}) \geq f(\mathbf{B} \cup \{\mathbf{x}\}) - f(\mathbf{B}), \quad \forall \mathbf{A} \subset \mathbf{B} \in 2^{\mathcal{X}_C}, \mathbf{x} \in \mathcal{X}_C \setminus \mathbf{B};$  (3)
- (c)  $f(\mathbf{A} \cup \{\mathbf{x}\}) - f(\mathbf{A}) \geq f(\mathbf{A} \cup \{\mathbf{x}, \mathbf{y}\}) - f(\mathbf{A} \cup \{\mathbf{y}\}), \quad \forall \mathbf{A}, \mathbf{B} \in 2^{\mathcal{X}_C}, \mathbf{x}, \mathbf{y} \in \mathcal{X}_C \setminus \mathbf{A}.$

Inequality (3) is known as the *diminishing returns* property, stating that the increment resulting of the addition of an element  $\mathbf{x}$  to a set is a decreasing function of the set to which it is added.

We say that a set-function  $f$  is non-decreasing when  $f(\mathbf{A} \cup \{\mathbf{x}\}) \geq f(\mathbf{A})$  for any  $\mathbf{A} \subset \mathcal{X}_C$  and any  $\mathbf{x} \in \mathcal{X}_C$ .

The problem of maximising a set-function is in general NP-hard, its exact resolution requiring the evaluation of  $f$  over all the  $2^C$  elements of the power set, which is infeasible except in trivial cases of little practical interest. In this paper we concentrate on variants of the greedy (one-step-ahead) algorithm (see below) to efficiently find approximate maximizers of size  $k$  of a given set-function  $f$ .

When several solutions exist at Step 3, a single one is selected (e.g. randomly). In spite of its simplicity, under mild conditions on  $f$  Algorithm 1 can be fairly efficient when applied to a non-decreasing submodular function as stated by the following theorem [1, 21].

**Theorem 1** *Let  $f$  be a non-decreasing submodular function, then, for any given  $k$ ,  $1 \leq k \leq C$ , Algorithm 1 returns a set  $\mathbf{X}$  with bounded optimality gap*

$$\frac{f^* - f(\mathbf{X})}{f^* - f(\emptyset)} \leq (1 - 1/k)^k \leq 1/e < 0.3679, \quad (4)$$

---

**Algorithm 1** (Greedy Algorithm)

---

```
1: set  $\mathbf{X} = \emptyset$ 
2: while  $|\mathbf{X}| < k$  do
3:   find  $\mathbf{x}$  in  $\mathcal{X}_C$  such that  $f(\mathbf{X} \cup \{\mathbf{x}\})$  is maximal
4:    $\mathbf{X} \leftarrow \mathbf{X} \cup \{\mathbf{x}\}$ 
5: end while
6: return  $\mathbf{X}$ 
```

---

where  $f^* = \max_{\mathbf{X} \subset \mathcal{X}_C: |\mathbf{X}| \leq k} f(\mathbf{X})$  and  $e = \exp(1)$ .

Although this theorem is proven in both [21] and [1], we sketch below a simpler proof that clearly reveals how both properties of  $f$ , submodularity and monotonicity, are required to guarantee a bound on the optimality gap of the greedy solutions.

Proof. For any  $j = 1, \dots, k$ , let  $\mathbf{X}_j^* = \{\mathbf{x}_1^*, \dots, \mathbf{x}_j^*\}$  denote a  $j$ -point subset of  $\mathcal{X}_C$  such that  $f(\mathbf{X}_j^*) = \max_{\mathbf{X} \in \mathcal{X}_C: |\mathbf{X}|=j} f(\mathbf{X})$ , with  $\mathbf{X}^* = \mathbf{X}_k^*$ ,  $f^* = f(\mathbf{X}^*)$ , and let  $\mathbf{X}_j = [\mathbf{x}_1, \dots, \mathbf{x}_j]$  denote the  $j$ -point subset selected by the greedy algorithm (any of such subsets in case of a draw at Step 3 of Algorithm 1). Let  $\mathbf{X}_0 = \mathbf{X}_0^* = \emptyset$ . As  $f$  is non-decreasing, for any  $j = 0, \dots, k-1$ , we have

$$f(\mathbf{X}^*) \leq f(\mathbf{X}^* \cup \mathbf{X}_j) = f(\mathbf{X}_j) + \sum_{i=1}^k [f(\mathbf{X}_j \cup \mathbf{X}_i^*) - f(\mathbf{X}_j \cup \mathbf{X}_{i-1}^*)].$$

The submodularity of  $f$  implies  $f(\mathbf{X}_j \cup \mathbf{X}_i^*) - f(\mathbf{X}_j \cup \mathbf{X}_{i-1}^*) \leq f(\mathbf{X}_j \cup \{\mathbf{x}_i^*\}) - f(\mathbf{X}_j)$  for all  $i$ , and thus

$$f(\mathbf{X}^*) \leq f(\mathbf{X}_j) + \sum_{i=1}^k [f(\mathbf{X}_j \cup \{\mathbf{x}_i^*\}) - f(\mathbf{X}_j)].$$

The greedy algorithm selects  $\mathbf{x}_j$  which corresponds to the best possible choice; therefore, for all  $i$ ,  $f(\mathbf{X}_j \cup \{\mathbf{x}_i^*\}) \leq f(\mathbf{X}_j \cup \{\mathbf{x}_j\}) = f(\mathbf{X}_{j+1})$  and

$$f(\mathbf{X}^*) \leq f(\mathbf{X}_j) + k [f(\mathbf{X}_{j+1}) - f(\mathbf{X}_j)]. \quad (5)$$

This is equivalent to

$$f(\mathbf{X}^*) - f(\mathbf{X}_{j+1}) \leq \left(1 - \frac{1}{k}\right) [f(\mathbf{X}^*) - f(\mathbf{X}_j)],$$

which implies, when we let  $j$  go from 0 to  $k-1$ ,

$$f(\mathbf{X}^*) - f(\mathbf{X}_k) \leq \left(1 - \frac{1}{k}\right)^k [f(\mathbf{X}^*) - f(\emptyset)], \quad (6)$$

which concludes the proof. ■

Note that (6) implies the efficiency bound

$$\frac{f(\mathbf{X}_k) - f(\emptyset)}{f(\mathbf{X}^*) - f(\emptyset)} \geq 1 - \frac{1}{e} > 63.2\% \quad (7)$$

for all  $k$  and that (5) with  $j = k - 1$  gives the simple on-line bound

$$f(\mathbf{X}_k^*) \leq f(\mathbf{X}_k) + (k - 1) [f(\mathbf{X}_k) - f(\mathbf{X}_{k-1})], \quad k \geq 1.$$

Other on-line bounds can be found in [21], also for the case of functions that are not necessarily non-decreasing; see also [14, Sect. 4.4].

## 2.2 The lazy-greedy algorithm

At each iteration of Algorithm 1 the maximum of  $f(\mathbf{X} \cup \{\mathbf{x}\})$  over all  $\mathbf{x} \in \mathcal{X}_C \setminus \mathbf{X}$  must be found. As presented in [18], the submodularity of  $f$  can be further exploited to restrict actual evaluations of  $f$  to a proper subset of  $\mathcal{X}_C \setminus \mathbf{X}$ . Although the worst-case complexity of the modified algorithm is still  $\mathcal{O}(kC)$  when  $k = o(C)$ , important gains in computational complexity are observed in practice; see Figure 4 in Section 3.6 for an illustration.

Denote by  $\delta_{\mathbf{X}}(\mathbf{x}) = f(\mathbf{X} \cup \{\mathbf{x}\}) - f(\mathbf{X}) \geq 0$  the improvement of  $f$  when  $\mathbf{x}$  is added to  $\mathbf{X}$ , and let  $\mathbf{X}_n$  be the greedy solution at iteration  $n$ . Also, remark that maximization of  $f(\mathbf{X}_n \cup \{\mathbf{x}\})$  in Step 3 of Algorithm 1 is equivalent to maximization of  $\delta_{\mathbf{X}_n}(\mathbf{x})$ .

By construction,  $\mathbf{X}_i \subset \mathbf{X}_n$  for all  $i < n$ . Then, since  $f$  is submodular, for all  $i < n$ ,  $\delta_{\mathbf{X}_n}(\mathbf{x}) \leq \delta_{\mathbf{X}_i}(\mathbf{x})$ ; i.e., for each  $\mathbf{x}$  the increments  $\delta_{\mathbf{X}_i}(\mathbf{x})$  decrease from iteration to iteration, as the size of  $\mathbf{X}_i$  grows. At the first iteration, we compute all  $\delta_{\mathbf{X}_0}(\mathbf{x})$  for all  $\mathbf{x} \in \mathcal{X}_C$ , establishing for each  $\mathbf{x}$  an upper bound  $\bar{\delta}(\mathbf{x})$  on  $\delta_{\mathbf{X}_k}(\mathbf{x})$  at subsequent iterations. These upper bounds are updated as follows.

Consider, at iteration  $k$ , scanning of the set  $\mathcal{X}_C \setminus \mathbf{X}_{k-1}$  to compute the solution of Step 3 of Algorithm 1. Let  $\mathcal{L}_{k-1} \subset \mathcal{X}_C$  denote the set of points that are possible solutions of Step 3. Initialize  $\mathcal{L}_{k-1} = \mathcal{X}_C \setminus \mathbf{X}_{k-1}$ , and let  $\mathbf{x}_k^{**}$  be its member with largest  $\bar{\delta}(\mathbf{x})$ . While  $\mathcal{L}_{k-1} \neq \emptyset$ , update  $\bar{\delta}(\mathbf{x}_k^{**}) = \delta_{\mathbf{X}_{k-1}}(\mathbf{x}_k^{**})$  and remove from  $\mathcal{L}_{k-1}$  all  $\mathbf{x}$  whose upper bound is smaller than or equal to  $\bar{\delta}(\mathbf{x}_k^{**})$ . When  $\mathcal{L}_{k-1} = \emptyset$ , update  $\mathbf{X}_{k-1}$  into  $\mathbf{X}_k = \mathbf{X}_{k-1} \cup \{\mathbf{x}_k^{**}\}$ .

This modified algorithm trades memory for computational power, requiring the storage of the most recently updated values of  $\delta_{\mathbf{X}}(\mathbf{x})$  for all  $\mathbf{x} \in \mathcal{X}_C \setminus \mathbf{X}$ .

## 3 Covering measures

### 3.1 Definitions

For any positive scalar  $r$ , we define the covering measure  $\Phi_r(\mathbf{Z}_n)$  of  $\mathbf{Z}_n$  by

$$\Phi_r(\mathbf{Z}_n) = \frac{\text{vol}\{\mathcal{X} \cap [\cup_{i=1}^n \mathcal{B}(\mathbf{z}_i, r)]\}}{\text{vol}(\mathcal{X})},$$

with  $\mathcal{B}(\mathbf{z}, r)$  the closed ball with center  $\mathbf{z}$  and radius  $r$ . For a given  $\mathbf{Z}_n$ , consider also the function  $r \in \mathbb{R}^+ \rightarrow F_{\mathbf{Z}_n}(r) = \Phi_r(\mathbf{Z}_n)$ .  $F_{\mathbf{Z}_n}$  is non-decreasing and satisfies  $F_{\mathbf{Z}_n}(0) = 0$  and  $F_{\mathbf{Z}_n}(r) = 1$  for any  $r \geq \text{CR}(\mathbf{Z}_n)$ . If  $X$  is distributed with the uniform probability measure  $\mu$  on  $\mathcal{X}$ , we have

$$\text{Prob}\{X \in \cup_{i=1}^n \mathcal{B}(\mathbf{z}_i, r)\} = \text{Prob}\{d(X, \mathbf{Z}_n) \leq r\} = \int_{\{\mathbf{x} \in \mathcal{X} : d(\mathbf{x}, \mathbf{Z}_n) \leq r\}} \mu(d\mathbf{x}) = F_{\mathbf{Z}_n}(r), \quad (8)$$

and  $F_{\mathbf{Z}_n}$  is the cumulative distribution function (c.d.f.) of the random variable  $d(X, \mathbf{Z}_n)$ , supported on  $[0, \text{CR}(\mathbf{Z}_n)]$ .



The set function  $\mathbf{Z}_n \subset \mathcal{X} \rightarrow \Phi_r(\mathbf{Z}_n)$  is non-decreasing and satisfies  $\Phi_r(\emptyset) = 0$ . Moreover, for any  $\mathbf{x} \in \mathcal{X}$ , the difference  $\Phi_r(\mathbf{Z}_n \cup \{\mathbf{x}\}) - \Phi_r(\mathbf{Z}_n)$  is non-increasing with respect to  $\mathbf{Z}_n$ , so that  $\Phi_r$  is submodular. In [24], the authors define criteria based on the distribution of distances to design points, in particular the  $(1 - \gamma)$ -covering radius  $r_{1-\gamma}(\mathbf{Z}_n)$  of  $\mathbf{Z}_n$  which corresponds to the  $(1 - \gamma)$ -quantile of  $d(X, \mathbf{Z}_n)$ . Using approximations valid for large  $d$  (see Remark 3), they investigate the properties of  $r_{1-\gamma}(\mathbf{Z}_n)$  for different random designs, in particular designs formed by  $n$  points i.i.d. in  $\mathcal{C}_d = [0, 1]^d$  with a beta distribution. In [23], they consider full-factorial and  $2^{d-1}$  fractional factorial designs, also for large  $d$ . We follow a different route, considering the values of  $F_{\mathbf{Z}_n}(r)$  taken at different  $r$  and exploiting the submodularity of  $\Phi_r$  to construct deterministic incremental designs  $\mathbf{X}_n$  which, thanks to Theorem 1, have guaranteed efficiency for all  $n$  for the criterion considered. Our constructions are effective also for small  $d$ , as several examples will illustrate.

### 3.2 A covering-based submodular criterion

Denote by  $f_{\mathbf{Z}_n}$  the probability density function (p.d.f.) corresponding to  $F_{\mathbf{Z}_n}$ . For any  $b, B$ ,  $0 \leq b < B$ ,  $q > -1$  and  $\mathbf{Z}_n \neq \emptyset$ , define the integrated covering measure

$$I_{b,B,q}(\mathbf{Z}_n) = \int_b^B r^q F_{\mathbf{Z}_n}(r) dr \quad (9)$$

$$= \frac{1}{q+1} \left\{ [B^{q+1} F_{\mathbf{Z}_n}(B) - b^{q+1} F_{\mathbf{Z}_n}(b)] - \int_b^B r^{q+1} f_{\mathbf{Z}_n}(r) dr \right\}, \quad (10)$$

with  $I_{b,B,q}(\mathbf{Z}_n) = 0$  for  $\mathbf{Z}_n = \emptyset$ . The set function  $I_{b,B,q} : \mathbf{Z}_n \rightarrow I_{b,B,q}(\mathbf{Z}_n)$  is non-decreasing and submodular, and satisfies  $I_{b,B,q}(\emptyset) = 0$ .

Since  $F_{\mathbf{Z}_n}(B) = 1$  for any  $B \geq \text{CR}(\mathbf{Z}_n)$ , maximizing  $I_{0,B,q}(\mathbf{Z}_n)$  with respect to  $\mathbf{Z}_n$  for  $B \geq \text{diam}(\mathcal{X})$  is equivalent to minimizing  $\int_0^B r^{q+1} f_{\mathbf{Z}_n}(r) dr = \mathbb{E}_n\{R^{q+1}\}$  in (10), where the random variable  $R$  has the p.d.f.  $f_{\mathbf{Z}_n}$  and where

$$(\mathbb{E}_n\{R^{q+1}\})^{1/(q+1)} = E_{q+1}(\mathbf{Z}_n) \quad (11)$$

is the  $L^{q+1}$ -mean quantization error induced by  $\mathbf{Z}_n$ , see [9]. It satisfies  $E_{q+1}(\mathbf{Z}_n) \leq \text{CR}(\mathbf{Z}_n)$  and  $E_{q+1}(\mathbf{Z}_n) \rightarrow \text{CR}(\mathbf{Z}_n)$  as  $q \rightarrow \infty$ . For  $B$  and  $q$  large enough, maximizing  $I_{0,B,q}(\mathbf{Z}_n)$  should therefore provide designs with small values of  $\text{CR}(\mathbf{Z}_n)$ ; moreover, the greedy maximization of  $I_{0,B,q}(\mathbf{Z}_n)$  is equivalent to the greedy minimization of  $E_{q+1}(\mathbf{Z}_n)$ , which is proved in [17] to ensure that  $E_{q+1}(\mathbf{Z}_n)$  tends to zero at rate  $n^{-1/d}$ . Now, since  $I_{b,B,q}$  is submodular, its greedy maximization with Algorithm 1 provides sequences of incremental designs  $\mathbf{X}_n$  with a  $I_{b,B,q}$ -efficiency at least 63.2% for any  $n$ , see (7). These observations motivate the investigations on the properties of  $I_{b,B,q}$  and its numerical evaluation presented in the paper.

**Remark 1 (Greedy minimization of the  $L^{q+1}$ -mean quantization error)** *When  $B \geq \text{diam}(\mathcal{X})$  ( $B \geq \text{diam}(\mathcal{X})/2$  when  $\mathcal{X}$  is convex and  $\mathbf{x}_1$  is its Chebyshev center), the greedy minimization of  $I_{0,B,q}(\mathbf{Z}_n)$ ,  $q > -1$ , selects*

$$\mathbf{x}_{n+1} \in \text{Arg max}_{\mathbf{x} \in \mathcal{X}} \int_{\mathcal{C}(\mathbf{x})} (d^{q+1}(\mathbf{z}, \mathbf{X}_n) - \|\mathbf{x} - \mathbf{z}\|^{q+1}) \mu(d\mathbf{z})$$

*at iteration  $n$ , with  $\mathcal{C}(\mathbf{x})$  the cell with generator  $\mathbf{x}$  in the Voronoi partition of  $\mathcal{X}$  associated with  $\mathbf{X}_n \cup \{\mathbf{x}\}$ .*



### 3.3 Choice of $b$ and $B$ for $\mathcal{X} = \mathcal{C}_d$

Taking  $b > 0$  will allow us to focus attention at each iteration on how well the design covers points far enough from its elements. In practice, it has moderate influence on performance but enables some simplification of the computations for small  $d$ , see Sections 3.4, 3.5 and 3.8.

Large enough values of  $B$  have no influence on  $I_{b,B,q}(\mathbf{Z}_n)$  as  $I_{b,B,q}(\mathbf{Z}_n) = I_{b,\text{CR}(\mathbf{Z}_n),q}(\mathbf{Z}_n)$  for any  $B \geq \text{CR}(\mathbf{Z}_n)$ . In particular, when  $B \geq \sqrt{d}/2$  then  $F_{\mathbf{Z}_n}(B) = 1$  for any  $\mathbf{Z}_n$  containing  $\mathbf{x}_1 = (1/2, \dots, 1/2)$ . The values of  $\text{CR}(\mathbf{X}_n)$  for incremental designs  $\mathbf{X}_n$  constructed with smaller values of  $B$  tend to decrease more smoothly and, the criterion being more localized,  $\mathbf{X}_n$  itself tends to fill the space more evenly; see Figures 5 and 6 for an illustration.

Supposing that the objective is to obtain a sequence of designs with good space-filling properties for  $n \in \{n_1, n_1 + 1, \dots, n_2\}$ , we can relate the choices of  $b$  and  $B$  to lower and upper bounds on  $\text{CR}_n^*$ , the optimal (minimum) value of  $\text{CR}(\mathbf{Z}_n)$ , for designs of size  $n_1$  and  $n_2$ .

Since the  $n$  balls  $\mathcal{B}(\mathbf{x}_i, \text{CR}_n^*)$  centered at the optimal design points cover  $\mathcal{X}$ ,  $nV_d(\text{CR}_n^*)^d \geq \text{vol}(\mathcal{X}) = 1$ , with  $V_d$  the volume of the  $d$ -dimensional unit ball  $\mathcal{B}(\mathbf{0}, 1)$ ,  $V_d = \pi^{d/2}/\Gamma(d/2 + 1)$ . This implies the following lower bound:

$$\text{CR}_n^* \geq R_*(n, d) = (nV_d)^{-1/d}. \quad (12)$$

We may note that, for  $n$  fixed,  $R_*(n, d) = \sqrt{d}/\sqrt{2\pi e} + \mathcal{O}(1/\sqrt{d})$ ,  $d \rightarrow \infty$ , showing that for large  $d$   $R_*(n, d)$  is not sensitive to  $n$ .

Conversely,  $\mathcal{X}$  can be covered by  $m^d$  hypercubes with side  $1/m$ , and therefore by  $m^d$  balls with radius  $\sqrt{d}/(2m)$ . Taking  $m = \lfloor n^{1/d} \rfloor$ , the largest integer smaller than or equal to  $n^{1/d}$ , we have  $n \geq m^d$  and therefore the following upper bound follows:

$$\text{CR}_n^* \leq \text{CR}_{m^d}^* \leq R^*(n, d) = \frac{\sqrt{d}}{2 \lfloor n^{1/d} \rfloor}. \quad (13)$$

A reasonable choice when  $n \in [n_1, n_2]$  is then  $b = b_* = R_*(n_2, d)$  and  $B = B^* = R^*(n_1, d)$ . Besides a simplification of calculations for the approximation of  $I_{b,B,q}(\mathbf{Z}_n)$ , see Sections 3.4 and 3.5, an effect of truncation of the integral (10) to a smaller value  $B^*$  is the existence of several equivalent solutions for  $\mathbf{x}$  at Step 3 of Algorithm 1. (However, as shown below, for moderate values of  $n$ , this truncation to  $B^*$  is only effective when  $d$  is small.)

The left panel of Figure 1 shows the bounds  $R_*(n, d)$  and  $R^*(n, d)$  as functions of  $n = 1, \dots, 250$  for  $d = 2$  (magenta  $\nabla$ ),  $d = 5$  (red  $\blackstar$ ) and  $d = 10$  (black  $\circ$ ). We have  $R^*(1, d) = \sqrt{d}/2$  for all  $d$ ; the  $k$ th jump downwards of  $R^*(n, d)$  occurs at  $n = (k + 1)^d$ . For example, the effect of choosing  $B = B^*$  instead of  $B = \sqrt{d}/2$  will be only be effective after  $n = 32$  when  $d = 5$  (see the left panel of Figure 8) and  $n = 1024$  for  $n = 10$ .

**Remark 2 (Initialization)** *When one is not interested in the performance of the smallest designs, with  $n \leq n_1$ , it seems preferable to directly choose the  $n_1$  first points by minimizing  $\text{CR}(\mathbf{Z}_{n_1})$ , and then initialize Algorithm 1 at this  $n_1$ -point design. The batch optimization of  $\text{CR}(\mathbf{Z}_{n_1})$  is a difficult task, but several methods available are able to produce designs with reasonably good performance; see, e.g., [25]. The investigation of properties and performance of incremental constructions initialized in this way is not considered in the present paper, where we consider the entire nested sequence, starting at  $n = 1$ . This will be the subject of future studies, together with backwards constructions that start with an optimal  $n_2$ -point design and iteratively eliminate design points until an  $n_1$ -point design is reached, for some  $n_1 < n_2$ .*

### 3.4 Evaluation of $I_{b,B,q}(\mathbf{Z}_n)$

To evaluate  $I_{b,B,q}(\mathbf{Z}_n)$ , we substitute the empirical c.d.f.  $\widehat{F}_{\mathbf{Z}_n}$  for  $F_{\mathbf{Z}_n}(r)$  in (9), where  $\widehat{F}_{\mathbf{Z}_n}$  is obtained by replacing  $\mu$  in (8) by the uniform measure  $\mu_Q$  supported on a finite subset  $\mathcal{X}_Q$  of  $\mathcal{X}$ ,

$$\mathcal{X}_Q = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(Q)}\}.$$

$\mathcal{X}_Q$  must be well spread over  $\mathcal{X}$ ; it may be a regular grid, or correspond to the first  $Q$  points of a low-discrepancy sequence in  $\mathcal{X}$ . This amounts to replacing  $F_{\mathbf{Z}_n}$  by the empirical c.d.f.  $\widehat{F}_{\mathbf{Z}_n}$ , based on the  $Q$  distances  $d_j(\mathbf{Z}_n) = d(\mathbf{x}^{(j)}, \mathbf{Z}_n)$ ,  $j = 1, \dots, Q$ , in (9). Denoting by  $\bar{d}_j$  the truncated version of  $d_j$ ,  $\bar{d}_j(\mathbf{Z}_n) = \min\{\max\{d_j(\mathbf{Z}_n), b\}, B\}$ ,  $j = 1, \dots, Q$ , and by  $\bar{d}_{j:Q}$  the  $\bar{d}_j$  sorted by increasing values, with  $\bar{d}_{1:Q}(\mathbf{Z}_n) \leq \bar{d}_{2:Q}(\mathbf{Z}_n) \leq \dots \leq \bar{d}_{Q-1:Q}(\mathbf{Z}_n) \leq \bar{d}_{Q:Q}(\mathbf{Z}_n) = B$ , we obtain

$$I_{b,B,q}(\mathbf{Z}_n) \approx \widehat{I}_{b,B,q}^A(\mathbf{Z}_n) = \frac{1}{Q(q+1)} \sum_{j=1}^Q j \left[ \bar{d}_{j+1:Q}^{q+1}(\mathbf{Z}_n) - \bar{d}_{j:Q}^{q+1}(\mathbf{Z}_n) \right]. \quad (14)$$

One may notice that (10) implies that, when  $b = 0$  and  $B$  is large enough (in particular, for any  $B \geq \text{diam}\mathcal{X}$ ), the maximization of  $\widehat{I}_{b,B,q}^A(\mathbf{Z}_n)$  is equivalent to the center location problem [18] defined by the minimization of  $\sum_i d_j^{q+1}(\mathbf{Z}_n)$ . As already mentioned in Section 3.2, it also corresponds to the minimization of the  $L^{q+1}$ -mean quantization error induced by  $\mathbf{Z}_n$ , which, in this discrete setting where  $\mu_Q$  is substituted for  $\mu$ , can be performed by clustering algorithms, see [15], including `kmeans` and variants based of the Chebyshev centers of the Voronoi cells defined by the elements of  $\mathbf{Z}_n$  [25].

An alternative approach is to evaluate  $I_{b,B,q}(\mathbf{Z}_n)$  through a discrete approximation of the integral in (9), with  $\widehat{F}_{\mathbf{Z}_n}$  substituted for  $F_{\mathbf{Z}_n}$ . Taking  $m$  radii  $r_i$  well spread over  $[b, B]$ , with  $b = r_1 < r_2 < \dots < r_m = B$ , the trapezoidal rule gives

$$I_{b,B,q}(\mathbf{Z}_n) \approx \widehat{I}_{b,B,q}^B(\mathbf{Z}_n) = \sum_{i=1}^{m-1} \frac{r_{i+1}^q \widehat{F}_{\mathbf{Z}_n}(r_{i+1}) + r_i^q \widehat{F}_{\mathbf{Z}_n}(r_i)}{2} (r_{i+1} - r_i). \quad (15)$$

When the  $r_i$  are regularly spaced, with  $r_{i+1} - r_i = \delta_r$  for all  $i$ , the expression simplifies into

$$\widehat{I}_{b,B,q}^B(\mathbf{Z}_n) = \delta_r \left[ \frac{r_1^q \widehat{F}_{\mathbf{Z}_n}(r_1) + r_m^q \widehat{F}_{\mathbf{Z}_n}(r_m)}{2} + \sum_{i=2}^{m-1} r_i^q \widehat{F}_{\mathbf{Z}_n}(r_i) \right].$$

The right panel of Figure 1 shows a plot of  $\widehat{F}_{\mathbf{Z}_n}(r_i)$  at 50 points regularly spaced in  $[b, B] = [0.05, 0.3]$  for two designs  $\mathbf{Z}_5 \subset \mathbf{Z}_{15}$  in  $\mathcal{C}_2$ . As expected, the curve for  $n = 15$  dominates the one for  $n = 5$ .

**Remark 3 (Large  $d$  approximation)** *Using an Edgeworth expansion of the Central Limit Theorem, in [36] the authors derive approximations of  $F_{\mathbf{x}}(r)$  that are valid for large  $d$ , with  $F_{\mathbf{x}}$  the c.d.f. for the one-point design  $\{\mathbf{x}\}$ . For  $X$  distributed with the uniform probability measure on  $\mathcal{X}$  and  $\mathbf{Z}_n$  a random design with  $n$  i.i.d. elements  $\mathbf{z}_i$ , we have  $F_{\mathbf{Z}_n}(r) = \text{Prob}\{d(X, \mathbf{Z}_n) \leq r\} = 1 - \prod_{i=1}^n [1 - F_{\mathbf{z}_i}(r)]$ ; this is also exploited in [24]. However, the approximation of  $F_{\mathbf{x}}$  is not directly exploitable for the estimation of  $I_{b,B,q}(\mathbf{Z}_n)$  for a given  $\mathbf{Z}_n$  since  $F_{\mathbf{Z}_n}(r) \neq 1 - \prod_{i=1}^n [1 - F_{\mathbf{z}_i}(r)]$ .*

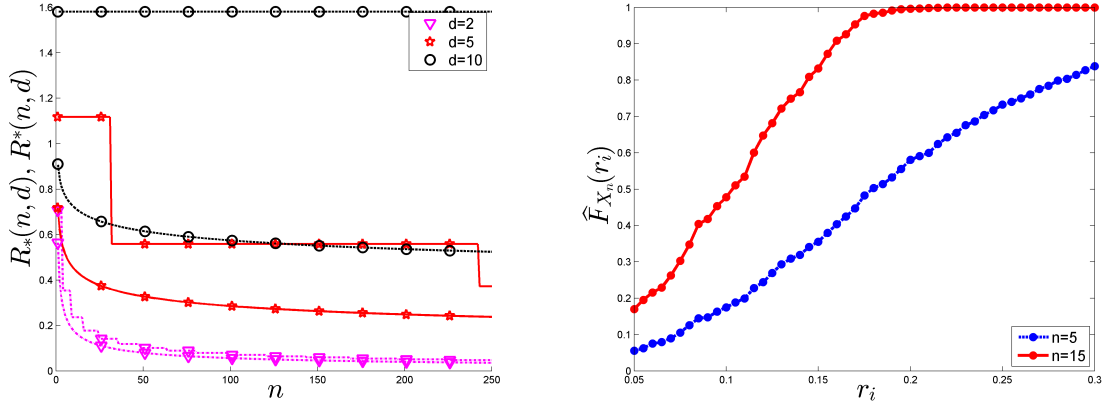


Figure 1: Left: values of  $R_*(n, d)$  and  $R^*(n, d)$ , respectively given by (12) and (13), for  $n = 1, \dots, 250$  and  $d = 2, 5, 10$ . Right: Empirical c.d.f.  $\hat{F}_{\mathbf{Z}_n}(r_i)$  at  $r_i = b + (i - 1)B/(m - 1)$ ,  $i = 1, \dots, m$ ,  $b = 0.05$  and  $B = 0.3$ , for  $n = 5, 15$  with  $\mathbf{Z}_5 \subset \mathbf{Z}_{15}$ .

### 3.5 Greedy maximization of $\hat{I}_{b,B,q}(\mathbf{Z}_n)$ with Algorithm 1

Denote by  $\mathcal{X}_C$  the set of candidate points in Algorithm 1 (with possibly  $\mathcal{X}_C = \mathcal{X}_Q$ , the support of the measure used to evaluate  $I_{b,B,q}$ , but this is not mandatory). For any design  $\mathbf{X}_n$  in the incremental construction, denote by  $\mathbf{D}_n$  the  $Q \times C$  matrix with elements  $\{\mathbf{D}_n\}_{i,j} = d_i(\mathbf{X}_n \cup \{\mathbf{x}^{(j)}\})$ ,  $\mathbf{x}^{(i)} \in \mathcal{X}_C$ . It can be computed recursively as follows. At initialization of Algorithm 1,  $\{\mathbf{D}_0\}_{i,j}$  is the  $Q \times C$  matrix of inter-distances between the points of  $\mathcal{X}_Q$  and those of  $\mathcal{X}_C$ . Then, when  $\mathbf{X}_{n+1} = \mathbf{X}_n \cup \{\mathbf{x}^{(k)}\}$  for some  $\mathbf{x}^{(k)} \in \mathcal{X}_C$ , each column  $\{\mathbf{D}_{n+1}\}_{\cdot,\ell}$  of  $\mathbf{D}_{n+1}$  is given by  $\min\{\{\mathbf{D}_n\}_{\cdot,\ell}, \{\mathbf{D}_0\}_{\cdot,k}\}$ , where the minimum is taken element-wise, for  $\ell = 1, \dots, C$ .

The elements of column  $\ell$  of  $\mathbf{D}_n$ , truncated to  $b$  from below and  $B$  from above, provide the value of the empirical c.d.f.  $\hat{F}_{\mathbf{X}_n}$  at each of the  $m$  specified  $r_i$  in  $[b, B]$  and yield the approximation  $\hat{I}_{b,B,q}^B(\mathbf{X}_n \cup \{\mathbf{x}^{(\ell)}\})$  of  $I_{b,B,q}(\mathbf{X}_n \cup \{\mathbf{x}^{(\ell)}\})$ , see (15), which is used to select the best  $\mathbf{x}^{(\ell)}$  in  $\mathcal{X}_C$ . The summation over the elements of column  $\ell$  of  $\mathbf{D}_n$ , truncated to the interval  $[b, B]$  and sorted by increasing values, gives the approximation  $\hat{I}_{b,B,q}^A(\mathbf{X}_n \cup \{\mathbf{x}^{(\ell)}\})$ , see (14).

### 3.6 Example 1: the greedy and lazy greedy algorithms

Here  $\mathcal{X}$  is the  $d$ -dimensional hypercube  $\mathcal{C}_d$  and the set  $\mathcal{X}_C$  of candidate points coincides with  $\mathcal{X}_Q$  which corresponds to the first  $Q$  elements of Sobol' sequence in  $\mathcal{X}$ .

We first compare the computational times<sup>1</sup> of the greedy and lazy-greedy maximizations<sup>2</sup> of  $\hat{I}_{0,\sqrt{d}/2,5}^A(\mathbf{Z}_n)$  and  $\hat{I}_{0,\sqrt{d}/2,5}^B(\mathbf{Z}_n)$  for various  $n$  and  $d = 2, 10$  with  $Q = 2^{11} = 2048$ . Figure 2 shows the tremendous acceleration provided by the lazy-greedy implementation and the linear increase of computational cost with  $n$  for the greedy version. For the lazy-greedy algorithm, maximization of  $\hat{I}_{0,\sqrt{d}/2,5}^B(\mathbf{Z}_n)$  (with  $m = 100 \ll Q$ ) is faster than maximization of  $\hat{I}_{0,\sqrt{d}/2,5}^A(\mathbf{Z}_n)$ ; see also Table 1 in Section 3.8. The effect of  $d$  is negligible.

<sup>1</sup>All calculations are made with Matlab, on a PC with a clock speed of 3.3 GHz and 8 GB RAM, and operations are vectorized whenever possible.

<sup>2</sup>Note that the designs obtained in the two cases are identical.

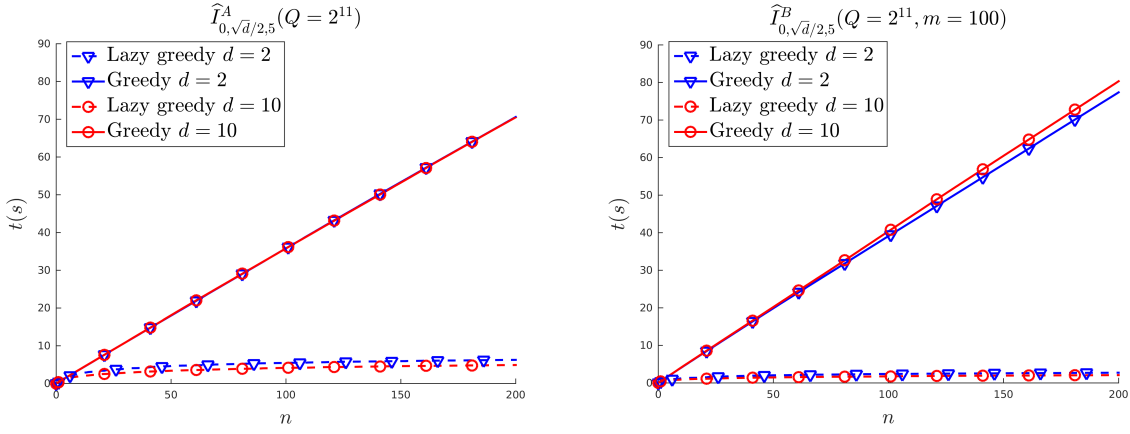


Figure 2: Computational times of the greedy and lazy-greedy maximizations of  $\widehat{I}_{0,\sqrt{d}/2,5}^A(\mathbf{Z}_n)$  and  $\widehat{I}_{0,\sqrt{d}/2,5}^B(\mathbf{Z}_n)$ .

Figure 3 shows the computational times of the maximizations of  $\widehat{I}_{0,\sqrt{d}/2,5}^A(\mathbf{Z}_n)$  and  $\widehat{I}_{0,\sqrt{d}/2,5}^B(\mathbf{Z}_n)$  (with  $m = 100$ ) for  $Q = 2^{10}$ ,  $2^{11}$  and  $2^{12}$  when  $d = 5$  and  $n = 150$ . The complexity of the greedy maximization of  $\widehat{I}_{0,\sqrt{d}/2,5}^A(\mathbf{Z}_n)$  can be shown to be dominated by the sorting operations, and is of order  $\mathcal{O}(nCQ \log Q)$  (here,  $\mathcal{O}(nQ^2 \log Q)$ ). It is of order  $\mathcal{O}(nmCQ)$  (here,  $\mathcal{O}(nmQ^2)$ ) for  $\widehat{I}_{0,\sqrt{d}/2,5}^B(\mathbf{Z}_n)$ . Since our Matlab code is vectorized, the left panel of Figure 3 does not reflect these dependencies on  $Q$  exactly. The observed time complexities for the lazy versions are coherent with this analysis, with important gains that are difficult to predict. We can understand the lazy algorithm as inducing a decrease of the effective size of the candidate set  $\mathcal{X}_C$  at each iteration, from  $C$  to  $\alpha C$ ,  $\alpha < 1$ . Denote by  $m_k$  the number of updates of  $\mathbf{x}_k^{**}$  at iteration  $k$  (with therefore  $m_1 = C$ ), see Section 2.2, and let  $\alpha_k = m_k/C$ : the effective size of  $\mathcal{X}_C$  for the lazy-greedy algorithm is  $m_k = C \alpha_k$  at iteration  $k$  and  $\bar{\alpha}_n C = (C/n) \sum_{k=1}^n \alpha_k$  in  $n$  iterations (i.e., for a  $n$ -point design). Figure 4 shows  $\alpha_k$  as a function of  $k$  for the maximization of  $\widehat{I}_{0,\sqrt{d}/2,5}^B(\mathbf{Z}_n)$  when  $Q = 2^{11}$ , with an average of  $\bar{\alpha}_n \simeq 0.063$  for  $n = 150$ . The value of  $\bar{\alpha}_{150}$  is rather stable: it belongs to the interval  $[0.062, 0.0655]$  for the three values of  $Q$  considered when we maximize  $\widehat{I}_{0,\sqrt{d}/2,5}^A(\mathbf{Z}_n)$  or  $\widehat{I}_{0,\sqrt{d}/2,5}^B(\mathbf{Z}_n)$ . If the cost of processing the  $C \times Q$  distances  $d(\mathbf{x}^{(i)}, \mathbf{X}_k \cup \{\mathbf{x}^{(j)}\})$ , for  $\mathbf{x}^{(i)} \in \mathcal{X}_Q$ ,  $\mathbf{x}^{(j)} \in \mathcal{X}_C$ , were  $C$  times the cost of processing the  $Q$  distances  $d(\mathbf{x}^{(i)}, \mathbf{X}_k \cup \{\mathbf{x}^{(j)}\})$  for one  $\mathbf{x}^{(j)} \in \mathcal{X}_C$ , the values of  $\bar{\alpha}_n$  would roughly coincide with the ratios of computing times displayed in Figure 3-right, but vectorization of operations in Matlab makes the acceleration of the lazy-greedy implementation smaller than expected. Moreover, as the size of the objects manipulated exceeds some threshold, the interplay between memory management and computational operations imposes overheads that are difficult to evaluate.

In the rest of the paper we shall only consider the lazy-greedy maximizations of  $\widehat{I}_{b,B,q}^A(\mathbf{Z}_n)$  and  $\widehat{I}_{b,B,q}^B(\mathbf{Z}_n)$ .

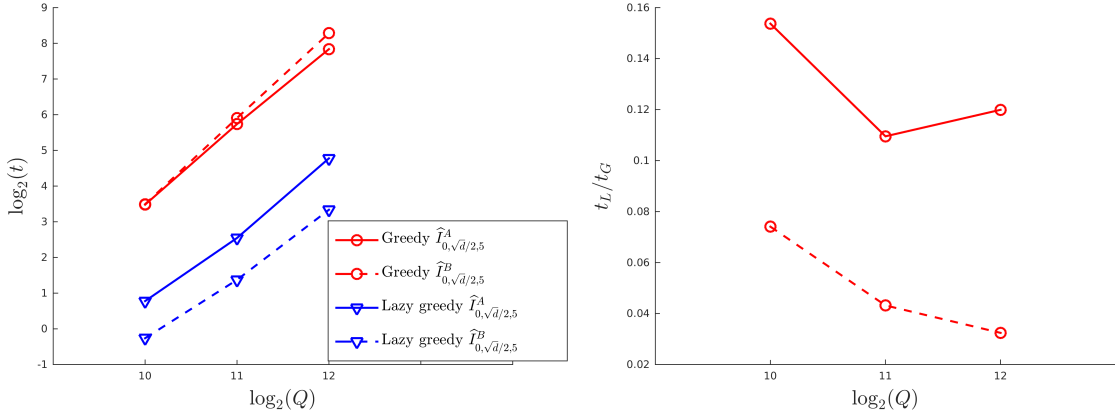


Figure 3: Left: computational times of the greedy and lazy-greedy maximizations of  $\hat{I}_{0,\sqrt{d}/2,5}^A(\mathbf{Z}_n)$  and  $\hat{I}_{0,\sqrt{d}/2,5}^B(\mathbf{Z}_n)$  (with  $m = 100$ ) for  $d = 5$ ,  $n = 150$  and  $Q = 2^{10}$ ,  $2^{11}$  and  $2^{12}$ . Right: ratio between computational times of the lazy-greedy and greedy versions (solid line for  $\hat{I}_{0,\sqrt{d}/2,5}^A(\mathbf{Z}_n)$  and dashed line for  $\hat{I}_{0,\sqrt{d}/2,5}^B(\mathbf{Z}_n)$ ).

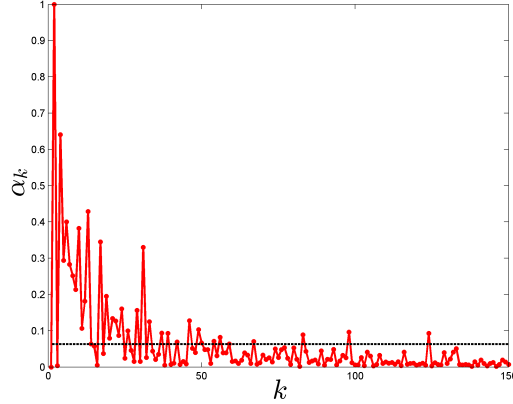


Figure 4: Effective size  $\alpha_k$  of  $\mathcal{X}_C$  as a function of  $k$  for the lazy-greedy maximization of  $\hat{I}_{0,\sqrt{d}/2,5}^B(\mathbf{Z}_n)$ , with  $m = 100$ ,  $d = 5$  and  $Q = 2^{11}$ ; the horizontal dashed line indicates the average value  $\bar{\alpha}_{150}$ .

### 3.7 Example 2: effect of $b, B$ and $q$

We apply Algorithm 1 to the maximization of the approximation  $\hat{I}_{b,B,q}^A(\mathbf{Z}_n)$  given by (14) when  $\mathcal{X}_C = \mathcal{X}_Q$  corresponds to the first  $2^{10} = 1024$  elements of Sobol' sequence in  $\mathcal{X} = \mathcal{C}_2$ .

We first compare the two cases  $b = 0$ ,  $B = \sqrt{d}/2$  and  $b = R_\star(n_2, d)$ ,  $B = B^\star(n_1, d)$  for  $q = 5$ , with  $n_1 = 10$  and  $n_2 = 20$ . Figure 5 presents the exact value of  $\text{CR}(\mathbf{X}_n)$  for  $n$  between 1 and 20 (red solid curve with  $\star$ ) and its under approximation  $\text{CR}_{\mathcal{X}_Q}(\mathbf{X}_n)$  given by (1) (black dashed

curve with  $\circ$ ). The blue dashed curve (+) presents the estimated<sup>3</sup> values of  $\text{CR}_n^*$ . The magenta curve ( $\nabla$ ) shows the empirical value of  $E_{q+1}(\mathbf{X}_n)$ , see (11) in Section 3.2, obtained for the uniform measure  $\mu_Q$  on  $\mathcal{X}_Q$ . Although on both panels the value  $q = 5$  is too small for  $E_{q+1}(\mathbf{X}_n)$  to provide a good approximation of  $\text{CR}(\mathbf{X}_n)$ , Algorithm 1 manages to ensure a reasonable decrease of  $\text{CR}(\mathbf{X}_n)$  along iterations. This decrease is much more regular when  $b = R_*(n_2, d)$  and  $B = B^*(n_1, d)$  than when  $b = 0$  and  $B = \sqrt{d}/2$ . Figure 6 shows the incremental designs that are constructed and helps understanding the different behaviors of  $\text{CR}(\mathbf{X}_n)$ : when  $b = 0$ ,  $B = \sqrt{d}/2$ ,  $\mathbf{x}_1$  corresponds to the optimal one-point design and is at the center of  $\mathcal{C}_2$ , the next two points  $\mathbf{x}_2$  and  $\mathbf{x}_3$  are close to those for the optimal 3-point design  $\{\mathbf{x}_1, \widehat{\mathbf{x}}_2, \widehat{\mathbf{x}}_3\}$  with  $\mathbf{x}_1$  fixed at  $(1/2, 1/2)$ ; see the Appendix. The situation is different when  $b = R_*(n_2, d)$  and  $B = B^*(n_1, d)$ , and we see on the right panel of Figure 6 that sacrificing optimality at a given  $n$  may be at the benefit of a more regular decrease of  $\text{CR}(\mathbf{X}_n)$ : here the choice of the initial points  $\mathbf{x}_1$  and  $\mathbf{x}_2$  is much worse than with  $b = 0$  and  $B = \sqrt{d}/2$  (with  $\text{CR}(\mathbf{X}_1) = 0.9326$  and  $\text{CR}(\mathbf{X}_2) = 0.7184$ ), but the situation quickly improves, leading in particular to much better values for  $n = 4, \dots, 8$ .

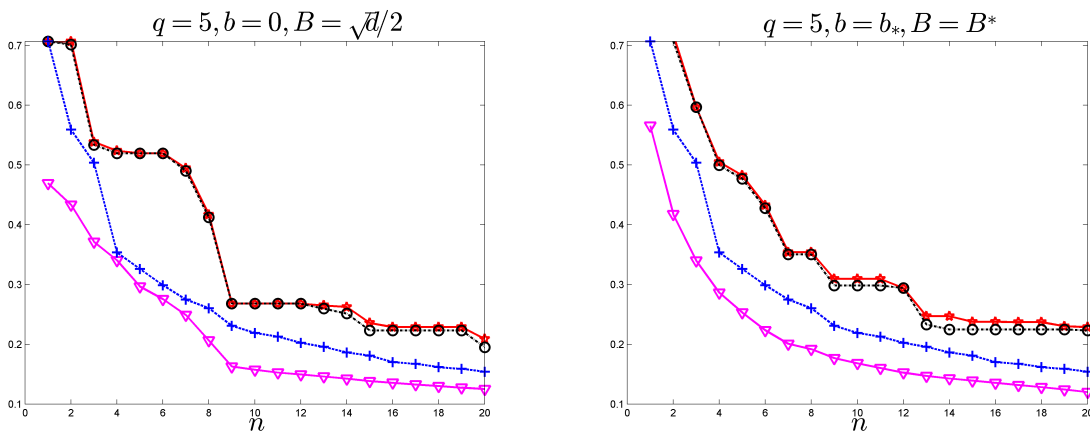


Figure 5:  $\text{CR}(\mathbf{X}_n)$  (red solid line and  $\star$ ),  $\text{CR}_{\mathcal{X}_Q}(\mathbf{X}_n)$  (black dashed curve and  $\circ$ ),  $\text{CR}_n^*$  (blue dashed curve and +), empirical value of  $E_{q+1}(\mathbf{X}_n)$  (11) (magenta curve and  $\nabla$ ), for  $\mathbf{X}_n$  obtained by greedy maximization of  $\widehat{I}_{b,B,q}^A(\mathbf{X}_n)$ .

We now consider the effect of  $q$ . Figure 7 concerns the same situation as on the right panels of Figure 5 and 6, but for  $q = 50$  instead of  $q = 5$ . The approximation  $E_{q+1}(\mathbf{X}_n)$  is closer to  $\text{CR}(\mathbf{X}_n)$  than in Figure 5, but  $\text{CR}(\mathbf{X}_n)$  is smaller on Figure 5-right for  $n \in \{1, 2, 3, 7, \dots, 12, 14, \dots, 20\}$ , indicating that best performances are not necessarily achieved with high values of  $q$ . This is directly visible when comparing the two designs  $\mathbf{X}_{20}$ , where big holes are present when  $q = 50$ .

### 3.8 Example 3: approximations $\widehat{I}_{b,B,q}^A$ and $\widehat{I}_{b,B,q}^B$

We compare the designs obtained by maximizing  $\widehat{I}_{b,B,q}^A(\mathbf{Z}_n)$  given by (14) and  $\widehat{I}_{b,B,q}^B(\mathbf{Z}_n)$  given by (15), with  $m = 100$ , in terms of  $\text{CR}(\mathbf{Z}_n)$ . As in Example 1, the set  $\mathcal{X}_C$  of candidate points

<sup>3</sup>Exact for  $n = 1, 2$ , best (smallest) values obtained for each  $n$  between 3 and 20 by running a `kmeans`-type clustering algorithm on a  $50 \times 50$  regular grid in  $\mathcal{X}$ ; see [25]. The values plotted are therefore not necessarily equal to the true values of  $\text{CR}_n^*$ , but we believe that the overestimation is negligible.

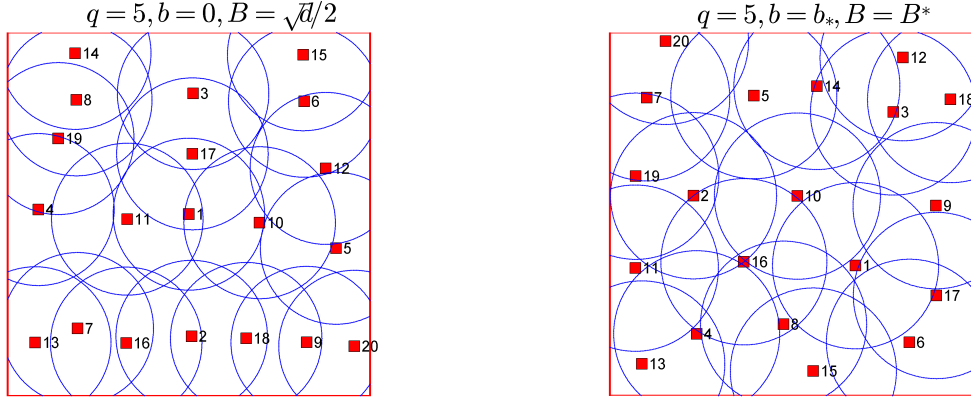


Figure 6:  $\mathbf{X}_{20}$  and circles centered at design points with radius  $\text{CR}(\mathbf{X}_{20})$ ; the order of selection of the points is indicated.

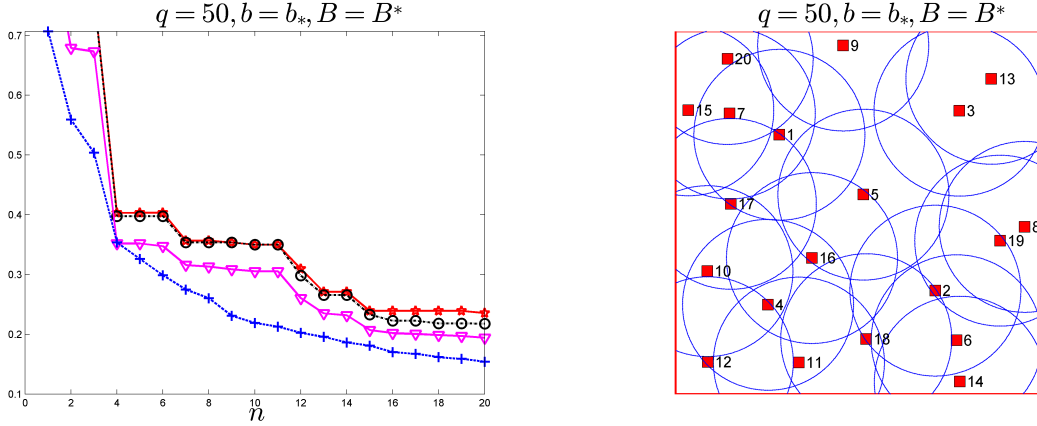


Figure 7: Left:  $\text{CR}(\mathbf{X}_n)$  (red solid line and  $\star$ ),  $\text{CR}_{\mathcal{X}_Q}(\mathbf{X}_n)$  (black dashed curve and  $\circ$ ),  $\text{CR}_n^*$  (blue dashed curve and  $+$ ), empirical value of  $E_{q+1}(\mathbf{X}_n)$  (11) (magenta curve and  $\nabla$ ), for  $\mathbf{X}_n$  obtained by greedy maximization of  $\widehat{I}_{b,B,q}^A(\mathbf{X}_n)$ . Right:  $\mathbf{X}_{20}$  and circles centered at design points with radius  $\text{CR}(\mathbf{X}_{20})$ ; the order of selection of the points is indicated.

coincides with  $\mathcal{X}_Q$  and corresponds to the first  $2^{11} = 2048$  elements of Sobol' sequence in  $\mathcal{X} = \mathcal{C}_d$ ; we consider the two cases  $d = 5$  and  $d = 10$  and take  $q = 5$ . The covering radius  $\text{CR}(\mathbf{Z}_n)$  is approximated by  $\text{CR}_{\mathcal{X}_N}(\mathbf{Z}_n)$  given by (1), with  $\mathcal{X}_N$  given by  $2^{18}$  points of a scrambled Sobol' sequence complemented by a  $2^d$  full factorial design, which gives  $N = 262176$  for  $d = 5$  and  $N = 263168$  for  $d = 10$ ; see the Appendix.

Figure 8 shows  $n^{1/d} \text{CR}_{\mathcal{X}_N}(\mathbf{X}_n)$  as a function of  $n$ , for  $d = 5$  and  $d = 10$ . We consider the two cases where  $b = b_\star = R_\star(n_2, d)$  and  $B = B^\star = R^\star(n_1, d)$ , with  $n_1 = 50$  and  $n_2 = 100$  and  $b = 0, B = \sqrt{d}/2$  (both cases yield  $\mathbf{x}_1 = (1/2, \dots, 1/2)$ ). The red  $\star$  are for  $\widehat{I}_{b_\star, B^\star, q}^A$ , the blue  $\nabla$  for  $\widehat{I}_{b_\star, B^\star, q}^B$ ; the magenta  $\circ$  are for  $\widehat{I}_{0, \sqrt{d}/2, q}^A$  and the black  $\times$  for  $\widehat{I}_{0, \sqrt{d}/2, q}^B$ .



There is no clear winner between  $\widehat{I}_{b^*,B^*,q}^A$  and  $\widehat{I}_{0,\sqrt{d}/2,q}^A$  when  $d = 5$ ; we merely note that for small  $n$  the fluctuations of  $\text{CR}(\mathbf{X}_n)$  are stronger with  $\widehat{I}_{0,\sqrt{d}/2,q}^A$ . The approximation  $\widehat{I}_{0,\sqrt{d}/2,q}^B(\mathbf{X}_n)$  for  $n \geq 32$  uses few values of  $r_i$  in the range of interest and is thus less accurate than  $\widehat{I}_{b^*,B^*,q}^B$ , which may explain the larger values observed for  $\text{CR}(\mathbf{X}_n)$ .

When  $d = 10$ , the four sequences perform nearly similarly for all  $n$ , with a small preference for  $\widehat{I}_{b^*,B^*,q}^A$  and  $\widehat{I}_{0,\sqrt{d}/2,q}^A$ . Other numerical experiments, not presented here, with different  $d$  indicate that the performances are similar when we replace  $B^* = R^*(n_1, d)$  by  $\sqrt{d}/(2n_1^{1/d}) \leq R^*(n_1, d)$ , or when we let  $b$  and  $B$  change with  $n$ , as  $b = R_*(n, d)$  and  $B = \sqrt{d}/(2n^{1/d})$  — which precludes the use of the lazy-greedy algorithm.

Table 1 gives the computational times for the construction of  $\mathbf{X}_n$  with the lazy-greedy algorithm in each case. The effect of the truncation of the distances  $d(\mathbf{x}^{(i)}, \mathbf{X}_n)$  to  $[b_*, B^*]$  on computational time vanishes as  $d$  increases, in agreement with the left panel of Figure 1.

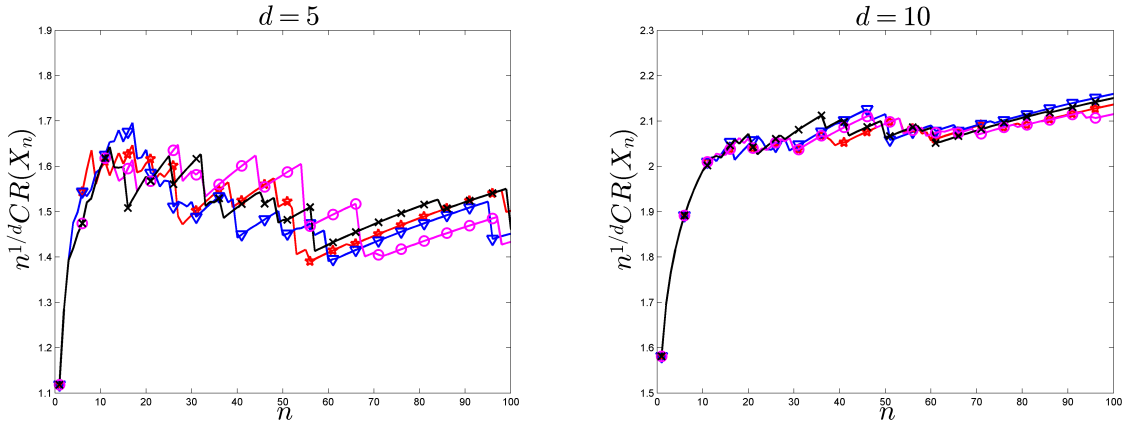


Figure 8:  $n^{1/d} \text{CR}_{\mathcal{X}_N}(\mathbf{X}_n)$  for  $n = 1, \dots, 100$  when maximizing  $\widehat{I}_{b^*,B^*,5}^A$  (red  $\star$ ),  $\widehat{I}_{b^*,B^*,5}^B$  (blue  $\nabla$ ),  $\widehat{I}_{0,\sqrt{d}/2,5}^A$  (magenta  $\circ$ ) and  $\widehat{I}_{0,\sqrt{d}/2,5}^B$  (black  $\times$ );  $m = 100$  for  $\widehat{I}_{b,B,q}^B$ .

Table 1: Computational times (in s) for the lazy-greedy algorithm in Example 3 (10 repetitions,  $n = 100$ ,  $C = Q = 2^{11}$ ,  $m = 100$  for  $\widehat{I}_{b,B,q}^B$ ).

	$\widehat{I}_{b^*,B^*,5}^A$	$\widehat{I}_{0,\sqrt{d}/2,5}^A$	$\widehat{I}_{b^*,B^*,5}^B$	$\widehat{I}_{0,\sqrt{d}/2,5}^B$
$d = 2$	2.4	5.3	0.8	2.3
$d = 5$	3.8	5.2	1.4	2.2
$d = 10$	4.1	4.1	1.8	1.8

## 4 Spacings and coffee-house design

In this section we present a second family of incremental algorithms, directly based on the most common geometrical criteria, the covering and packing radii. Even if these criteria are

not submodular, and thus Theorem 1 cannot be invoked to guarantee a lower bound on the efficiency of their greedy optimization, an efficiency of 50% can still be guaranteed.

Coffee-house designs [19], [20, Chap. 4] are obtained by greedy maximization of  $\text{PR}(\mathbf{Z}_n)$  for  $n > 1$ ; see [12] for an early suggestion. The first design point  $\mathbf{x}_1$  is usually chosen at the Chebyshev center of  $\mathcal{X}$  (the center of the minimal-radius ball enclosing  $\mathcal{X}$ ); then, at any  $n \geq 1$ ,  $\mathbf{x}_{n+1} \in \text{Arg max}_{\mathbf{x} \in \mathcal{X}} d(\mathbf{x}, \mathbf{X}_n)$ . Denote by  $\text{PR}_n^*$  the optimal (maximum) value of  $\text{PR}(\mathbf{Z}_n)$  for a  $n$ -point design in  $\mathcal{X}$ . The two ratios  $\text{CR}_n^*/\text{CR}(\mathbf{Z}_n)$  and  $\text{PR}(\mathbf{Z}_n)/\text{PR}_n^*$  are less than one by construction and quantify the efficiency of a design  $\mathbf{Z}_n$  in terms of its covering and packing radii. It is remarkable that the simple greedy coffee-house construction ensures the following property (irrespective of the choice made for  $\mathbf{x}_1$ ):

$$\frac{1}{2} \leq \frac{\text{CR}_n^*}{\text{CR}(\mathbf{X}_n)} \leq 1 \quad (n \geq 1) \quad \text{and} \quad \frac{1}{2} \leq \frac{\text{PR}(\mathbf{X}_n)}{\text{PR}_n^*} \leq 1 \quad (n \geq 2). \quad (16)$$

Indeed, by construction  $\text{PR}(\mathbf{X}_{n+1}) = d(\mathbf{x}_{n+1}, \mathbf{X}_n)/2 = \text{CR}(\mathbf{X}_n)/2$  for all  $n \geq 1$ . Take any  $n$ -point design  $\mathbf{Z}_n = \{\mathbf{z}_1, \dots, \mathbf{z}_n\}$  in  $\mathcal{X}$ . From the pigeonhole principle, one of the balls  $\mathcal{B}(\mathbf{z}_i, \text{CR}(\mathbf{Z}_n))$  must contain two points  $\mathbf{x}_i, \mathbf{x}_j$  of  $\mathbf{X}_{n+1}$ , implying that  $\text{PR}(\mathbf{X}_{n+1}) \leq \|\mathbf{x}_i - \mathbf{x}_j\|/2 \leq \text{CR}(\mathbf{Z}_n)$ . Therefore,  $\text{CR}_n^* \geq \text{PR}(\mathbf{X}_{n+1}) = \text{CR}(\mathbf{X}_n)/2$ . Similarly,  $\text{PR}_{n+1}^* \leq \text{CR}_n^* \leq \text{CR}(\mathbf{X}_n) = 2 \text{PR}(\mathbf{X}_{n+1})$ . The original proof is given in [8].

As for the maximization of  $I_{b,B,q}(\mathbf{Z}_n)$  in Section 3, the implementation of Algorithm 1 is much facilitated when  $\mathbf{x}_{n+1} \in \text{Arg max}_{\mathbf{x} \in \mathcal{X}_C} d(\mathbf{x}, \mathbf{X}_n)$  with  $\mathcal{X}_C$  a finite set of candidates. The efficiencies given in (16) remain valid provided that  $\text{CR}(\mathbf{X}_n)$  is approximated by  $\text{CR}_{\mathcal{X}_C}(\mathbf{X}_n)$  and that  $\text{CR}_n^*$  and  $\text{PR}_n^*$  are relative to optimal  $n$ -point designs in  $\mathcal{X}_C$ . Contrary to the maximization of  $\hat{I}_{b,B,q}(\mathbf{Z}_n)$  in Section 3.5, greedy maximization of  $\text{PR}(\mathbf{Z}_n)$  does not require computation the  $C \times C$  matrix of inter-distances between all points in  $\mathcal{X}_C$ , but only the update of the  $C$  distances  $d(\mathbf{x}^{(i)}, \mathbf{Z}_n)$  for the  $\mathbf{x}^{(i)}$  in  $\mathcal{X}_C$ . For that reason, the size  $C$  of the candidate set can be taken much larger than in Section 3.5.

The mesh-ratio  $\rho(\mathbf{Z}_n)$  is another interesting indicator of the quality of a design (to be minimized):

$$\rho(\mathbf{Z}_n) = \frac{\text{CR}(\mathbf{Z}_n)}{\text{PR}(\mathbf{Z}_n)} \quad (n \geq 2).$$

The inverse of  $\rho(\mathbf{Z}_n)$  is sometimes called uniformity measure; see, e.g., [3]. We shall denote  $\rho_{\mathcal{X}_N}(\mathbf{Z}_n)$  the mesh-ratio of  $\mathbf{Z}_n$  when  $\text{CR}(\mathbf{Z}_n)$  is approximated by  $\text{CR}_{\mathcal{X}_N}(\mathbf{Z}_n)$ , see (1). As a convex set cannot be covered by two or more non-overlapping balls having their centers in  $\mathcal{X}$ ,  $\text{PR}(\mathbf{Z}_n) < \text{CR}(\mathbf{Z}_n)$ , implying that  $\rho(\mathbf{Z}_n) > 1$ . When  $\mathbf{X}_n$  is a coffee-house design,  $\text{PR}(\mathbf{X}_{n+1}) = \text{CR}(\mathbf{X}_n)/2$  for  $n \geq 1$ , and since  $\text{CR}(\mathbf{Z}_n) \geq \text{CR}(\mathbf{Z}_{n+1})$  for all  $n$  and any nested designs  $\mathbf{Z}_n \subset \mathbf{Z}_{n+1}$ ,  $\rho(\mathbf{X}_n)$  satisfies

$$1 \leq \rho(\mathbf{X}_n) \leq 2 \quad (n \geq 2). \quad (17)$$

By construction, a coffee-house design places design points on the boundary of  $\mathcal{X}$ , which is not favourable to the objective of constructing designs with low values of  $\text{CR}(\mathbf{Z}_n)$ . For that reason, we consider below a modified method that forces design points to stay away from the boundary of  $\mathcal{X}$ .

Following [10]<sup>4</sup>, we define the maximal spacing  $S(\mathbf{Z}_n)$  as the radius of the largest ball contained in  $\mathcal{X}$  and not containing any of the  $\mathbf{z}_i$  in  $\mathbf{Z}_n$ ,

$$S(\mathbf{Z}_n) = \sup \{r : \exists \mathbf{x} \text{ such that } \mathbf{x} \oplus r \mathcal{B}(\mathbf{0}, 1) \subset \mathcal{X} \setminus \mathbf{Z}_n\},$$

with  $\oplus$  denoting the Minkowski sum. We slightly extend this notion by introducing a parameter  $\beta$  that controls the ratio between the distance to the design and the distance to the boundary  $\partial\mathcal{X}$  of the compact set  $\mathcal{X}$ , and define the  $\beta$ -spacing of  $\mathbf{Z}_n$ , for  $\beta > 0$ , as

$$S_\beta(\mathbf{Z}_n) = \sup \left\{ r : \exists \mathbf{x} \in \mathcal{X} \text{ such that } d(\mathbf{x}, \mathbf{Z}_n) \geq r \text{ and } d(\mathbf{x}, \partial\mathcal{X}) \geq \frac{r}{\beta} \right\} = \sup_{\mathbf{x} \in \mathcal{X}} D_\beta(\mathbf{x}, \mathbf{Z}_n),$$

where

$$D_\beta(\mathbf{x}, \mathbf{Z}_n) = \min \{d(\mathbf{x}, \mathbf{Z}_n), \beta d(\mathbf{x}, \partial\mathcal{X})\}, \quad \mathbf{x} \in \mathcal{X}, \quad (18)$$

and  $d(\mathbf{x}, \partial\mathcal{X}) = \inf_{\mathbf{z} \in \partial\mathcal{X}} \|\mathbf{x} - \mathbf{z}\|$ . Figure 9 gives an illustration for a 4-point design in the square  $[0, 1]^2$  with two different values of  $\beta$ . We have  $S_1(\mathbf{Z}_n) = S(\mathbf{Z}_n)$  and we define  $S_\infty(\mathbf{Z}_n) = \text{CR}(\mathbf{Z}_n)$ . We also define

$$P_\beta(\mathbf{Z}_n) = \min_{\mathbf{z}_i \neq \mathbf{z}_j \in \mathbf{Z}_n} \frac{1}{2} \min \{\|\mathbf{z}_i - \mathbf{z}_j\|, \beta d(\mathbf{z}_i, \partial\mathcal{X})\} \text{ and } \rho_\beta(\mathbf{Z}_n) = \frac{S_\beta(\mathbf{Z}_n)}{P_\beta(\mathbf{Z}_n)} \quad (n \geq 2), \quad (19)$$

and set  $P_\infty(\mathbf{Z}_n) = \text{PR}(\mathbf{Z}_n)$ , so that  $\rho_\infty(\mathbf{Z}_n) = \rho(\mathbf{Z}_n)$ .

The coffee-house algorithm can be straightforwardly extended to the greedy maximization of  $P_\beta(\mathbf{Z}_n)$ ,  $\beta > 0$ , using  $\mathbf{x}_{n+1} \in \text{Arg max}_{\mathbf{x} \in \mathcal{X}} D_\beta(\mathbf{x}, \mathbf{X}_n)$  for any  $n \geq 1$  (and if we use the convention  $\mathbf{X}_0 = \emptyset$  and  $d(\mathbf{x}, \emptyset) = +\infty$  for any  $\mathbf{x}$ , the same rule sets  $\mathbf{x}_1$  at the Chebyshev center of  $\mathcal{X}$ ). Following the same steps as for coffee-house design, we obtain the following property, whose proof is given in the Appendix.

**Theorem 2** *Let  $\mathcal{X}$  be a compact and convex subset of  $\mathbb{R}^d$  and let  $\mathbf{X}_n$  be a design obtained with the following greedy construction:  $\mathbf{X}_0 = \emptyset$  and  $\mathbf{x}_{k+1} \in \text{Arg max}_{\mathbf{x} \in \mathcal{X}} D_\beta(\mathbf{x}, \mathbf{X}_k)$ ,  $k = 0, 1, \dots, n-1$ , where  $D_\beta$  is defined in (18),  $\beta > 0$ , and  $d(\mathbf{x}, \emptyset) = +\infty$ . Then  $P_\beta(\mathbf{X}_{n+1}) = (1/2) S_\beta(\mathbf{X}_n)$  for  $n \geq 1$  and*

$$\frac{1}{2} \leq \frac{S_{\beta,n}^*}{S_\beta(\mathbf{X}_n)} \leq 1 \quad (n \geq 1) \quad \text{and} \quad \frac{1}{2} \leq \frac{P_\beta(\mathbf{X}_n)}{P_{\beta,n}^*} \leq 1, \quad 1 \leq \rho_\beta(\mathbf{X}_n) \leq 2 \quad (n \geq 2), \quad (20)$$

where  $S_{\beta,n}^* = \min_{\mathbf{Z}_n \subset \mathcal{X}} S_\beta(\mathbf{Z}_n)$  and  $P_{\beta,n}^* = \max_{\mathbf{Z}_n \subset \mathcal{X}} P_\beta(\mathbf{Z}_n)$ .

**Remark 4 (Lazy-greedy coffee-house design)** *Any nested designs  $\mathbf{Z}_n \subset \mathbf{Z}_{n+1}$  satisfy  $D_\beta(\mathbf{x}, \mathbf{Z}_{n+1}) \leq D_\beta(\mathbf{x}, \mathbf{Z}_n)$  for all  $\mathbf{x} \in \mathcal{X}$ . When a finite candidate set  $\mathcal{X}_C$  is substituted for  $\mathcal{X}$  in the greedy maximization of  $P_\beta(\mathbf{Z}_n)$ , this non-increasing property of  $D_\beta$  can be used to store a list of upper bounds  $\bar{D}_\beta(\mathbf{x}^{(i)})$  on  $D_\beta(\mathbf{x}^{(i)}, \mathbf{X}_k)$  for all  $\mathbf{x}^{(i)} \in \mathcal{X}_C$ , valid at any iteration  $k$ , initialized with  $\bar{D}_\beta(\mathbf{x}^{(i)}) = D_\beta(\mathbf{x}^{(i)}, \emptyset) = \beta d(\mathbf{x}^{(i)}, \partial\mathcal{X})$ . A lazy-greedy version of the algorithm is thus straightforward to implement; see Section 2.2. However, the simplicity of the updating of  $D_\beta(\mathbf{x}, \mathbf{X}_n)$  into  $D_\beta(\mathbf{x}, \mathbf{X}_n \cup \{\mathbf{x}_{n+1}\}) = \min\{D_\beta(\mathbf{x}, \mathbf{X}_n), \|\mathbf{x} - \mathbf{x}_{n+1}\|\}$  makes the lazy-greedy implementation less efficient than simple greedy that updates all  $D_\beta(\mathbf{x}^{(i)}, \mathbf{X}_n)$  at each iteration.*

<sup>4</sup>In [10],  $\mathcal{X}$  is only assumed to be bounded and bounded convex sets other than  $\mathcal{B}(\mathbf{0}, 1)$  are also considered.

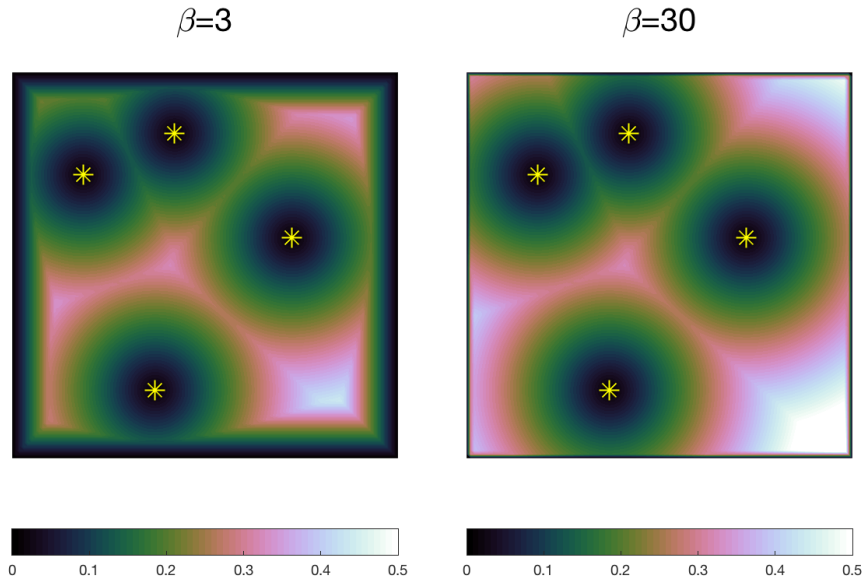


Figure 9: Edgephobe compromise distance  $D_\beta(\mathbf{x}, \mathbf{Z}_n)$  for a 4-point design (yellow stars) in  $\mathcal{X} = [0, 1]^2$ .

Algorithm 2 of [33], also based on geometrical considerations, corresponds to the greedy maximization of  $P_\beta(\mathbf{Z}_n)$  in  $\mathcal{X} = \mathcal{C}_d$ , for the value  $\beta = 2\sqrt{2d}$  (chosen in [33] by trial and error). Considering that the covering radius is of special importance, we recommend the following choice for  $\beta$  when  $\mathcal{X} = \mathcal{C}_d$ : since  $\mathbf{x}_1 = \mathbf{1}_d/2$ , the center of  $\mathcal{X}$ , when  $\beta = \infty$  the second point  $\mathbf{x}_2$  coincides with a vertex of  $\mathcal{X}$ . We suggest to choose  $\beta$  depending on the maximum target design size  $n_2$ , such that  $\mathbf{x}_2$  be at distance  $R_\star(n_2, d)$  from a vertex of  $\mathcal{X}$ , with  $R_\star(n_2, d)$  the lower bound on  $\text{CR}_{n_2}^\star$  given by (12) ( $R_\star(n_2, d)$  coincides with the value of  $b$  proposed in Section 3.3). This implies  $d(\mathbf{x}_2, \partial\mathcal{X}) = R_\star(n_2, d)/\sqrt{d} = \|\mathbf{x}_2 - \mathbf{x}_1\|/\beta$  and gives

$$\beta = \beta_\star(n_2, d) = \frac{d}{2R_\star(n_2, d)} - \sqrt{d}. \quad (21)$$

The left panel of Figure 16 of the Appendix shows  $\beta_\star(n, d)$  as a function of  $d$  for  $n = 50$  (red  $\star$ ),  $n = 100$  (blue  $\nabla$ ) and  $n = 200$  (black  $\circ$ ); the curve with magenta  $\times$  corresponds to the values suggested in [33]. We see that as  $n$  grows  $\beta_\star(n, d)$  increases (since we expect designs that better cover  $\mathcal{X}$ ), and as  $d$  becomes large the value of  $\beta_\star(n, d)$  exhibits a very slow growth after an initial fast decay (being almost constant for large values of  $n$ ).

## 5 Numerical study

This section illustrates the performance of the design algorithms presented in the previous sections, comparing among them and checking how well they compare to existing alternative

methods. Besides coffee-house design and its variants discussed in section 4, the study considers designs that are prefixes of low-discrepancy sequences (namely of the Sobol' and Halton sequences) and two incremental constructions proposed in [28] which we will briefly recall below.

Designs are compared both in terms of the covering radius  $\text{CR}(\mathbf{X}_n)$  as well as of the mesh-ratio  $\rho(\mathbf{X}_n)$ , which combines the values of the covering and packing radii. We stress that for the incremental design methods studied in this paper, the entire evolution of the performance indicators over a target range of design sizes is of interest, and not just their value for a given final design size. Since  $\text{CR}$  ideally decreases as  $n^{-1/d}$  all our plots will show its normalized version  $n^{1/d} \text{CR}(\mathbf{X}_n)$ , and an ideal design method should lead to covering-radius trajectories close to horizontal.

Two different situations are analysed in this section. In Example 4 the design space has a simple geometry embedded in a relatively large dimensional space,  $\mathcal{X} = \mathcal{C}_d = [0, 1]^{10}$ , the 10-dimensional unit hypercube; target design sizes go up to  $n = 10d$ , in accordance with the rule  $n = 10d$  (see [16] for a justification)<sup>5</sup>. Example 5 considers a non-convex (annular) domain  $\mathcal{X}$ . (In Example 4 we will also investigate how the geometry of the finite set of candidate points from which the design points are chosen impacts the quality of the designs defined, by letting  $\mathcal{X}_C$  be the elements of a Latin hypercube design of size  $n$ .)

We compare the performance of designs  $\mathbf{X}_n^A$  that incrementally maximize the covering measure  $\widehat{I}_{b^*, B^*, q}^A(\mathbf{X}_n)$  given by (14), to (i) nested subsequences  $\mathbf{X}_n^H$  and  $\mathbf{X}_n^S$  of low discrepancy Halton and Sobol' sequences, (ii)  $\mathbf{X}_n^{CH, \beta}$  obtained by greedy maximization of the criteria  $P_\beta(\mathbf{Z}_n)$  based on pairwise distances discussed in the previous section, see (19), and (iii) the two incremental constructions presented next (whose designs are denoted by  $\mathbf{X}_n^{VD}$  and  $\mathbf{X}_n^{RD}$ , see precise definitions below). The relative efficiency of the different design methods will be assessed through the performance-complexity tradeoff.

**Minimization of a relaxed version of  $\text{CR}(\mathbf{Z}_n)$ .** In [28], the following relaxed version of the coverage criterion of [30] is considered:

$$\Psi_q(\mathbf{Z}_n) = \Psi_q(\mathbf{Z}_n; \mu) = \left[ \int_{\mathcal{X}} \left( \frac{1}{n} \sum_{i=1}^n \|\mathbf{z}_i - \mathbf{x}\|^{-q} \right)^{-1} \mu(d\mathbf{x}) \right]^{1/q}, \quad q \neq 0, \quad (22)$$

with  $\mu$  the uniform probability measure on  $\mathcal{X}$ . For any  $n$ -point design  $\mathbf{Z}_n$ , it satisfies  $\Psi_q(\mathbf{Z}_n) \rightarrow \text{CR}(\mathbf{Z}_n)$  as  $q \rightarrow \infty$ . If  $\xi_n = (1/n) \sum_{i=1}^n \delta_{\mathbf{z}_i}$  denotes the empirical measure associated with  $\mathbf{Z}_n$ , we can define  $\psi_q(\xi_n) = \Psi_q(\mathbf{Z}_n)$  and more generally

$$\psi_q(\xi) = \psi_q(\xi; \mu) = \left[ \int_{\mathcal{X}} \left( \int_{\mathcal{X}} \|\mathbf{z} - \mathbf{x}\|^{-q} \xi(d\mathbf{z}) \right)^{-1} \mu(d\mathbf{x}) \right]^{1/q}, \quad q \neq 0,$$

for any probability measure  $\xi$  on  $\mathcal{X}$ . It is shown in [28] that  $\psi_q^q(\cdot)$  is a convex functional for  $q > 0$  (strictly convex for  $q \in (0, d)$ ); for reasons explained in that paper,  $q$  should be taken larger than  $\max\{0, d - 2\}$  and we use  $q = d$  in the examples below. The criterion  $\psi_q(\xi)$  can be minimized by the conditional (or constrained) gradient method of [7]; when the step

<sup>5</sup>Very small designs, with  $n \leq 10$ , are not meaningful and will be excluded from the analysis. Note that since  $n^{1/d} R_*(n, d) = V_d^{-1/d} \gtrsim 0.9106$  and  $n^{1/d} R^*(n, d) = \sqrt{d}/2 \lesssim 1.5812$  for  $n < 1024$ , which implies for the optimal (normalized) covering radius  $n^{1/d} \text{CR}_n^* \in (0.9106, 1.5812)$ .

size at iteration  $k$  equals  $1/(k+1)$ , the method corresponds to Wynn's Vertex-Direction (VD) method [35] of approximate design theory (in practice, a discrete measure  $\mu_Q$  supported on  $\mathcal{X}_Q \subset \mathcal{X}$  is substituted for  $\mu$ , and the minimization of  $\psi_q(\xi)$  corresponds to an  $A$ -optimal design problem). Taking the initial measure  $\xi^{(1)}$  as the one-point measure supported at  $\mathbf{z}_1 \in \mathcal{X}$  and using iteratively  $\xi^{(k+1)} = [k/(k+1)]\xi^{(k)} + \delta_{\mathbf{z}_{k+1}}/(k+1)$ , then, for each  $n \geq 1$ ,  $\xi^{(n)}$  has  $n$  support points which define an incremental design  $\mathbf{X}_n$ . Here,  $\mathbf{z}_1 \in \text{Arg min}_{\mathbf{z} \in \mathcal{X}} \int_{\mathcal{X}} \|\mathbf{z} - \mathbf{x}\|^q \mu(d\mathbf{x})$  and  $\mathbf{z}_{k+1}$  minimizes the directional derivative of  $\psi_q^q(\cdot)$  at  $\xi^{(k)}$  in the direction of the delta measure  $\delta_{\mathbf{z}}$  with respect to  $\mathbf{z}$ . This corresponds to

$$\mathbf{z}_{k+1} \in \text{Arg max}_{\mathbf{z} \in \mathcal{X}} \int_{\mathcal{X}} \left[ \|\mathbf{z} - \mathbf{x}\|^{-q} \left( \int_{\mathcal{X}} \|\mathbf{y} - \mathbf{x}\|^{-q} \xi^{(k)}(d\mathbf{y}) \right)^{-2} \right] \mu(d\mathbf{x});$$

see [28]. We shall denote by  $\mathbf{X}_n^{VD}$  the designs constructed in this manner, with  $\mu_Q$  substituted for  $\mu$ .

Straightforward calculations indicate that, for  $q > 0$ ,  $(1/n)\Psi_q^q(\mathbf{Z}_n, \mu_Q)$ , a discrete unnormalized form of  $\Psi_q^q(\mathbf{Z}_n; \mu)$  defined in (22), is non-increasing and supermodular (i.e.,  $-(1/n)\Psi_q^q(\mathbf{Z}_n, \mu_Q)$  is submodular). Theorem 1 thus applies to the greedy maximization of  $-(1/n)\Psi_q^q(\mathbf{Z}_n, \mu_Q)$ ; we shall denote by  $\mathbf{X}_n^{RD}$  a design obtained by greedy minimization of  $\Psi_q(\mathbf{Z}_n, \mu_Q)$ , a Relaxed and Discretized version of  $\text{CR}(\mathbf{Z}_n)$ . The lazy-greedy algorithm of Section 2.2 can also be applied, see Table 2.

**Example 4.** In this example  $\mathcal{X} = \mathcal{C}_{10}$ , and the following implementation choices were made. For the construction of  $\mathbf{X}_n^A$  we fix  $q = 5$  and  $b_\star = R_\star(100, d)$ ,  $B^\star = R^\star(50, d)$ , see (12), (13). Variants of coffee-house design  $\mathbf{X}_n^{CH, \beta}$  use  $\beta \in \{\beta_\star(n_2, d), 2\sqrt{2}, \infty\}$ , and for  $\mathbf{X}_n^{VD}$  and  $\mathbf{X}_n^{RD}$  we use  $q = 10$ . As in Example 3 (Section 3.8), the candidate set  $\mathcal{X}_C$  coincides with the set  $\mathcal{X}_Q$  used for  $\mu_Q$ , consisting of the first  $Q$  elements  $\mathbf{s}_i$  of Sobol' sequence in  $\mathcal{C}_{10}$ ;  $\text{CR}(\mathbf{Z}_n)$  is approximated by  $\text{CR}_{\mathcal{X}_N}(\mathbf{Z}_n)$  given by (1), with  $\mathcal{X}_N$  consisting of  $2^{18}$  points of a scrambled Sobol' sequence complemented by a  $2^d$  full factorial design ( $N = 263\,168$ ). For coffee-house designs we follow the suggestion in [33] and use a candidate set  $\mathcal{X}_{C'}$  given by the first  $C' = 1\,000d + 2n_2$  points of the Sobol' sequence. Although  $C' = 10\,200 > C = 4\,096$ , the constructions based on spacings are much faster than for  $\mathbf{X}_n^A$ , see Table 2. The constructions of  $\mathbf{X}_n^{VD}$  and  $\mathbf{X}_n^{RD}$  require that the candidate set  $\mathcal{X}_C$  does not intersect  $\mathcal{X}_Q$ ; for these designs, and following [28], we take the  $C = Q$  points  $\mathbf{s}_{4Q+1}, \dots, \mathbf{s}_{5Q}$  of Sobol' sequence in  $\mathcal{C}_{10}$ .

We start by comparing the incremental optimization of the integrated covering criterion with simple use of the prefixes of low-discrepancy sequences, through the covering radii and mesh-ratios of the resulting designs. The left panel of Figure 10 shows  $\text{CR}(\mathbf{X}_n)$ ,  $n = 10, \dots, 100$ , for  $\mathbf{X}_n^A$  (red  $\star$ ),  $\mathbf{X}_n^H$  (Halton sequence, blue  $\nabla$ ) and  $\mathbf{X}_n^S$  (Sobol' sequence, black  $\times$ ). This plot shows that the greedy maximization of  $\hat{I}_{b_\star, B^\star, q}^A(\mathbf{Z}_n)$  yields designs with significantly smaller covering radii than the two classical low discrepancy sequences, with the Sobol' sequence outperforming the Halton sequence. One may also notice that, due to the incremental construction,  $\text{CR}(\mathbf{X}_n^A)$  is significantly larger than the upper bound on  $\text{CR}_n^\star$  — equal to  $1.5812n^{-1/d}$  — the rate of decrease with  $n$  being correct, however.

The right panel of Figure 10 compares the covering radii of  $\mathbf{X}_n^A$  (same as on the left panel) to those of designs  $\mathbf{X}_n^{CH, \beta}$  produced by variants of the coffee-house design methodology. The following code is used:  $\beta = \infty$  (coffee-house design) in black  $\times$ ,  $\beta = 2\sqrt{2}d$  in magenta  $\circ$ , and  $\beta = \beta_\star(100, d)$  given by (21) in blue  $\nabla$ . We can see that  $\mathbf{X}_n^A$  has again the best overall performance,

only  $\mathbf{X}_n^{CH,\beta_*(100,d)}$  yielding smaller covering radii for a few design sizes over the range of values of  $n$  considered. Classic coffee-house design  $\mathbf{X}_n^{CH,\infty}$  is almost always outperformed by the two variants considered, in particular for the smaller design sizes. The relative merits of  $\mathbf{X}_n^{CH,2\sqrt{2}d}$  and  $\mathbf{X}_n^{CH,\beta_*(100,d)}$  are less clear, with the latter nested sequence of designs displaying, however, a more regular behavior as  $n$  varies<sup>6</sup>.

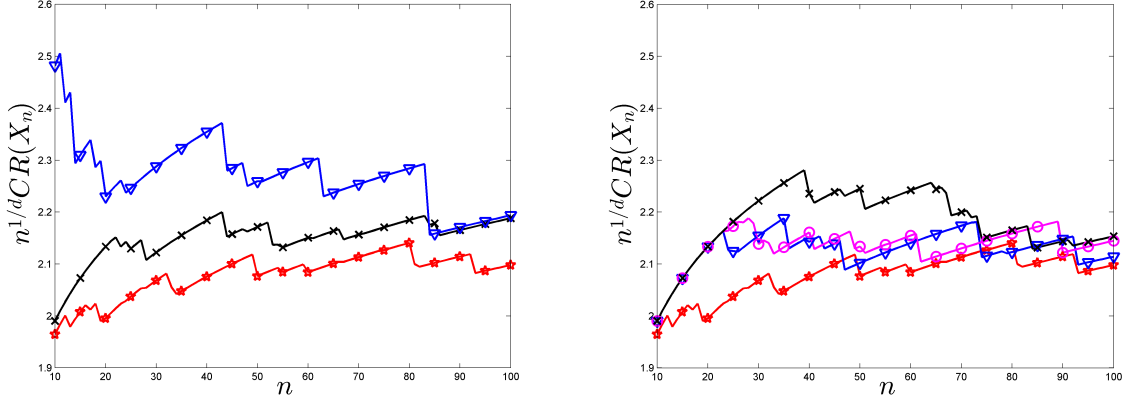


Figure 10:  $n^{1/d} \text{CR}_{\mathcal{X}_N}(\mathbf{X}_n)$ ,  $n = 10, \dots, 100$ ,  $d = 10$ . Left:  $\mathbf{X}_n^A$  (red  $\star$ ),  $\mathbf{X}_n^H$  (blue  $\nabla$ ) and  $\mathbf{X}_n^S$  Sobol' sequence (black  $\times$ ). Right:  $\mathbf{X}_n^A$  (red  $\star$ ),  $\mathbf{X}_n^{CH,\infty}$  (black  $\times$ ),  $\mathbf{X}_n^{CH,2\sqrt{2}d}$  (magenta  $\circ$ ) and  $\mathbf{X}_n^{CH,\beta_*(100,d)}$ , see (21), (blue  $\nabla$ ).

Figure 11 plots the values of the mesh-ratio  $\rho_{\mathcal{X}_N}(\mathbf{X}_n)$ , revealing a different picture: minimizing the covering radius tends to reduce the packing radius of a design, and  $\rho(\mathbf{X}_n)$  is very sensitive to the value of  $\text{PR}(\mathbf{X}_n)$ . We can see that  $\mathbf{X}_n^A$  (red  $\star$ ) has now worse performance (largest mesh-ratio) than all the other designs considered, being worst than the prefixes of low-discrepancy sequences (left panel), which in turn are also outperformed by coffee-house and its variants (right panel). The best mesh-ratios are observed for  $\mathbf{X}_n^{CH,\infty}$ , which greedily maximizes  $\text{PR}(\mathbf{X}_n)$ . Notice that  $\mathbf{X}_n^S$  (Sobol' sequence) performs better than  $\mathbf{X}_n^H$  (Halton sequence) also in terms of the mesh-ratio.

We assess now the impact of properties of the set of candidate points  $\mathcal{X}_C$ , by repeating the comparison between  $\mathbf{X}_n^A$  and the coffee-house designs  $\mathbf{X}_n^{CH,\beta}$  when the set of candidate points considered is  $\mathcal{X}_C \equiv \mathbf{Z}_{Lh,100}$ , a latin hypercube design. We hope in this way to define an order in  $\mathbf{Z}_{Lh,100}$  such that all nested designs  $\mathbf{Z}_{Lh,1:n} = \{\mathbf{z}_i\}_{i=1}^n$ ,  $\mathbf{z}_i \in \mathbf{Z}_{Lh,100}$ ,  $n \leq 100$ , have small covering radius<sup>7</sup>.

<sup>6</sup>Note that the comparisons in [33] show that the greedy maximization of  $P_\beta(\mathbf{Z}_n)$  with  $\beta = 2\sqrt{2}d$  performs better in terms of covering radius than sliced and nested Latin hypercube (Lh) designs proposed in the literature, which are only batch-incremental (see the references in [33]). A one-shot Lh design with maximum packing radius may nevertheless perform significantly better, both in terms of packing and covering radius. For instance, the covering radius of the 100-point Lh design for  $d = 10$  on the site <https://spacefillingdesigns.nl/> is  $\text{CR}_{\mathcal{X}_N}(\mathbf{X}_{Lh,100}) \simeq 1.2515$ , and  $100^{1/10} \text{CR}_{\mathcal{X}_N}(\mathbf{X}_{Lh,100}) \simeq 1.9836$ ; compare with the larger values achieved at the right end of the plots on Figure 10; see also Figure 12-left.

<sup>7</sup>When the points in  $\mathbf{Z}_{Lh,100}$  are arranged as in <https://spacefillingdesigns.nl/>, the order of selection of the 100 points for the design  $\mathbf{X}_n^A$  below is 65, 34, 94, 5, 100, 2, 8, 92, 22, 82, 29, 80, 90, 30, 40, 55, 31, 6, 33, 13, 83, 75,



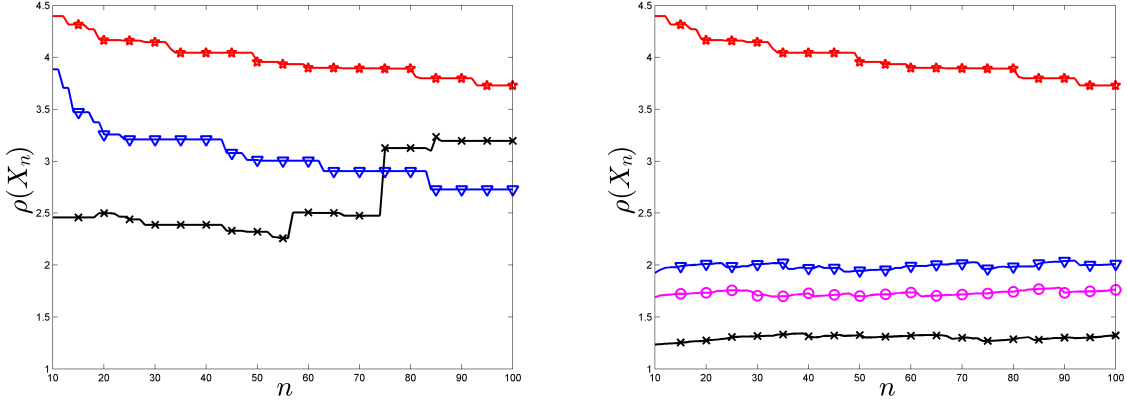


Figure 11:  $\rho_{\mathcal{X}_N}(\mathbf{X}_n)$  for  $n = 10, \dots, 100$ ,  $d = 10$ . Left:  $\mathbf{X}_n^A$  (red  $\star$ ),  $\mathbf{X}_n^H$  (blue  $\nabla$ ) and  $\mathbf{X}_n^S$  Sobol' sequence (black  $\times$ ). Right:  $\mathbf{X}_n^A$  (red  $\star$ ),  $\mathbf{X}_n^{CH,\infty}$  (black  $\times$ ),  $\mathbf{X}_n^{CH,2\sqrt{2}d}$  (magenta  $\circ$ ) and  $\mathbf{X}_n^{CH,\beta_*(100,d)}$ , see (21), (blue  $\nabla$ ).

While values of  $b$  and  $B$  that vary with  $n$  prevent use of lazy-greedy optimization, for a candidate set  $\mathcal{X}_C$  of such a small size ( $C = 100$ ) it is possible to implement the full greedy optimization of  $\hat{I}_{b,B,q}^A(\mathbf{Z}_n)$ . We use  $b = b_n = R_*(n, d)$ , see (12), and  $B = B_n = \sqrt{d}/(2n^{1/d})$ . Figure 12 shows the covering radii (left panel) and mesh-ratios (right panel) of designs  $\mathbf{X}_n^A$  along with those of  $\mathbf{X}_n^{CH,\beta}$  for  $\beta \in \{\infty, 2\sqrt{2}d, \beta_*(100, d)\}$ , obtained by greedily maximizing  $P_\beta(\mathbf{Z}_n)$  over the same set of candidate points. The figure shows that for this extremely constrained candidate set all methods yield similar overall performance trajectories, with the magenta and blue curves, corresponding to  $\beta = 2\sqrt{2}d$  and  $\beta = \beta_*(100, d)$ , being coincident. Note that coffee-house ( $\beta = \infty$ ) yields designs with a better mesh-ratio but with much larger covering radius when  $80 \leq n \leq 99$ . Comparing with the right panels in Figures 10 and 11 we can see that covering radii are now in general worse for the very small designs but better — except for  $\mathbf{X}_n^{CH,\infty}$  — for large design sizes (in particular the final value, for  $n = 100$ , is smaller than it was before), while the mesh-ratio of all methods is now close to the values achieved previously by the best method (classic coffee-house). This example suggests that incremental optimization of the criteria proposed in the paper produce, when applied to a final design with good space filling properties, sequences of nested designs with good overall space filling properties.

We finally compare the covering radii of  $\mathbf{X}_n^A$ , to the relaxed designs  $\mathbf{X}_n^{VD}$  and  $\mathbf{X}_n^{RD}$ , in the same situation as in Figure 10. Figure 13 presents the normalized covering radii for two values of  $Q \in \{2^{11}, 2^{12}\} = \{2048, 4096\}$ , showing that  $\mathbf{X}_n^A$  (red  $\star$ ) and  $\mathbf{X}_n^{RD}$  (black  $\times$ ) perform rather similarly, with a small advantage to  $\mathbf{X}_n^A$  — observed over the entire trajectory for the smaller value of  $Q$  — which is able to effectively exploit the denser candidate set to achieve smaller final covering radii for  $Q = 10^{12}$ . Designs  $\mathbf{X}_n^{VD}$  are more sensitive to  $Q$  (i.e., to the thinness of  $\mathcal{X}_Q$  and  $\mathcal{X}_C$ ) than the other two.

---

11, 59, 99, 16, 28, 76, 72, 89, 17, 93, 88, 1, 15, 61, 24, 81, 10, 85, 4, 71, 52, 86, 18, 68, 27, 79, 47, 64, 20, 19, 66, 58, 70, 77, 74, 51, 26, 98, 32, 63, 62, 46, 60, 69, 56, 43, 57, 25, 39, 87, 41, 73, 21, 48, 7, 96, 12, 37, 78, 91, 36, 9, 49, 42, 3, 35, 38, 23, 14, 53, 45, 95, 84, 50, 97, 54, 44, 67.

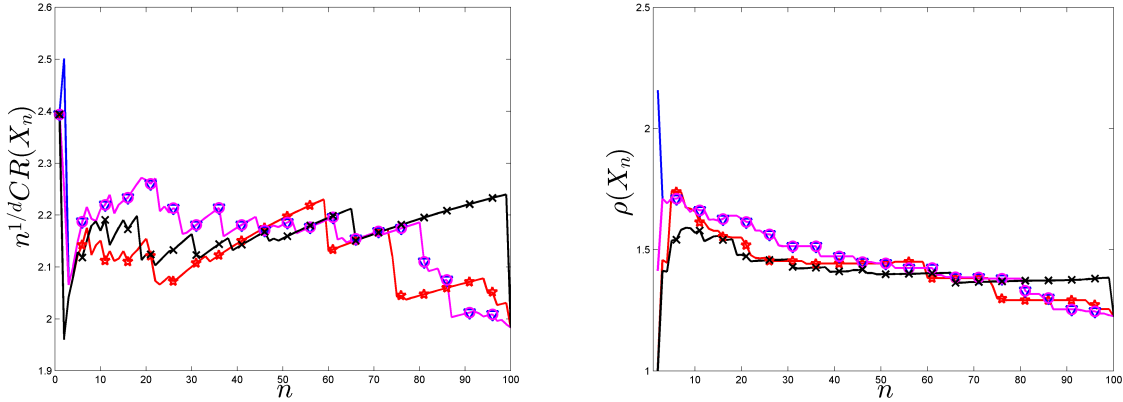


Figure 12: Left:  $n^{1/d} \text{CR}_{\mathcal{X}_N}(\mathbf{X}_n)$ , right:  $\rho_{\mathcal{X}_N}(\mathbf{X}_n)$ ;  $n = 1, \dots, 100$ ,  $d = 10$ : greedy maximization of  $\hat{I}_{b_n, B_n, 5}^A(\mathbf{Z}_n)$  (red  $\star$ ) and of  $P_\beta(\mathbf{Z}_n)$  with  $\beta = \infty$  (black  $\times$ ),  $\beta = 2\sqrt{2d}$  (magenta  $\circ$ ) and  $\beta = \beta_\star(n_2, d)$  given by (21) (blue  $\nabla$ ), all constructions use the candidate set  $\mathcal{X}_{100} = \mathbf{Z}_{Lh, 100}$ .

The material presented above suggests that  $\mathbf{X}_n^A$  may be an attractive alternative to classic incremental constructions: it clearly outperforms low discrepancy sequences, gives better overall performance than the best variants of coffee-house, and has a slight advantage, in particular in terms of stability, over the relaxed designs  $\mathbf{X}_n^{VD}$  and  $\mathbf{X}_n^{RD}$ . The computational load of the design methods compared is, however, significantly different. Table 2 shows the computational time of each design for the two values of  $Q$  considered, with  $\mathbf{X}_n^{LRD}$  the design obtained by lazy-greedy minimization of  $\Psi_q(\mathbf{Z}_n, \mu_Q)$  given by (22), showing that the good performance of  $\mathbf{X}_n^A$  comes at the price of increased execution times. The coffee-house variants are by far the fastest, closely followed by the vertex-direction method for the minimization of  $\psi_q(\xi; \mu_Q)$ . The other three methods have comparable computational times — the lazy-greedy version being only slightly faster than the simple greedy one for the minimization of  $\Psi_q(\mathbf{Z}_n, \mu_Q)$  due to the simplicity of the evaluation and updating of the criterion. We see thus that given its superior performance  $\mathbf{X}_n^A$  should be the preferred design method only when computational costs are not a prime concern. Whenever numerical complexity must be kept to a minimum, one of the variants of coffee-house design based on spacings will define designs with slightly increased covering radii at a much smaller cost. Note that any of these two alternative incremental design methods will yield designs with better space filling properties than simple use of prefixes of low-discrepancy sequences or even common greedy optimization of the packing radius (i.e., classic coffee-house).

Table 2: Computational time of  $\mathbf{X}_n$  (in s, 10 repetitions).

	$\mathbf{X}_n^A$	$\mathbf{X}_n^{VD}$	$\mathbf{X}_n^{RD}$	$\mathbf{X}_n^{LRD}$	$\mathbf{X}_n^{CH, \infty}$	$\mathbf{X}_n^{CH, 2\sqrt{2d}}$	$\mathbf{X}_n^{CH, \beta_\star(100, d)}$
$Q = 2^{11}$	4.1	0.7	5.6	4.4	0.3	0.6	0.6
$Q = 2^{12}$	15.8	2.8	21.4	17.6	0.3	0.6	0.6

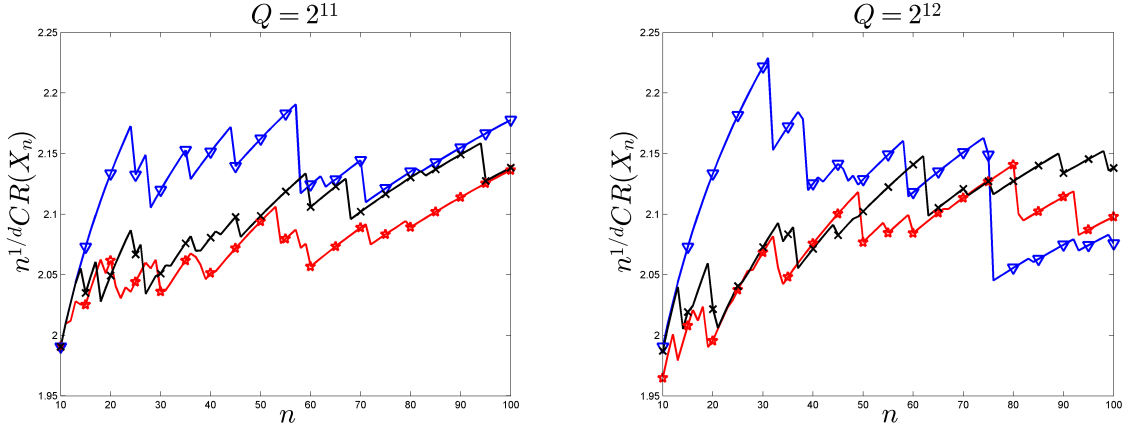


Figure 13:  $n^{1/d} \text{CR}_{\mathcal{X}_N}(\mathbf{X}_n)$ ,  $n = 10, \dots, 100$ ,  $d = 10$ , for  $\mathbf{X}_n^A$  (red  $\star$ ),  $\mathbf{X}_n^{VD}$  (blue  $\nabla$ ) and  $\mathbf{X}_n^{RD}$  (black  $\times$ ).

**Example 5.** We investigate in this example the impact of non-convex domains  $\mathcal{X}$ , by considering an annular geometry, letting  $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^2 : 1/2 \leq \|\mathbf{x}\| \leq 1\}$  (for simplicity, we still base the covering and packing radii on the Euclidean distance, although geodesics could have been used as well). Again, we will compare the maximization of the integrated covering criterion to use of prefixes of low-discrepancy sequences and to coffee-house design with  $\beta = \infty$ .

Implementation details are as follow. For  $\mathcal{X}_C$  we use the first  $C = 2048$  points of a Sobol' sequence of points in  $\mathcal{C}_2$  — renormalized to  $[-1, 1]^2$  — falling inside  $\mathcal{X}$  (with  $C$  close to  $1000d + 2n_2 = 2200$ , the value used in [33]). For the set used to approximate  $\mu$  on  $\mathcal{X}$  we use  $\mathcal{X}_Q \equiv \mathcal{X}_C$ . An analogous construction is used for the set  $\mathcal{X}_N$  used to approximate  $\text{CR}(\mathbf{Z}_n)$ , by retaining the first  $N = 2^{18}$  points of a renormalized scrambled Sobol' sequence that fall inside  $\mathcal{X}$ .

Designs  $\mathbf{X}_n^A$ , obtained by greedy maximization of  $\hat{I}_{b_\star, B_\star, 5}^A(\mathbf{Z}_n)$ , use  $b_\star = (\sqrt{3}\pi/2)R_\star(100, d)$  and  $B_\star = (\sqrt{3}\pi/2)R^\star(50, d)$ <sup>8</sup>. The low-discrepancy designs  $\mathbf{X}_n^S$  considered are the first  $n$  elements of  $\mathcal{X}_C$ .

Spacings are difficult to handle for non-convex domains, explaining why only  $\beta = \infty$  is studied in this example. Instead, we consider two distinct candidate sets in the construction of coffee-house designs:  $\mathcal{X}_C$ , the points of the Sobol' renormalized and clipped sequence, and, in an effort to enforce boundary avoidance, we will also use a different candidate set  $\mathcal{X}'_C$ , which is the result of restricting the renormalized Sobol' sequence to the eroded annulus  $\mathcal{X}' = \{\mathbf{x} \in \mathbb{R}^2 : 1/2 + r \leq \|\mathbf{x}\| \leq 1 - r\} \subset \mathcal{X}$ , with  $r = R_\star(100, d)/2$ .

Figure 14 shows the 100-point designs obtained. The left panel shows  $\mathbf{X}_n^A$  (red  $\blacksquare$ ) and  $\mathbf{X}_n^S$  (black dots), and the right panel displays coffee-house designs  $\mathbf{X}_n^{CH, \infty}$  (magenta  $\blacksquare$ ) and  $\mathbf{X}_n^{CH', \infty}$  (blue dots) obtained for candidate sets  $\mathcal{X}$  and  $\mathcal{X}'$ , respectively. Visually, they all fill  $\mathcal{X}$  reasonably well, although  $\mathbf{X}_n^S$  has a few nearly coincident points.

Figure 15 shows the performance of the four designs, in terms of normalized covering radius (left panel) and mesh-ratio (right panel). The three greedy constructions clearly outperform the Sobol' sequence (for which  $\rho_{\mathcal{X}_N}(\mathbf{X}_n^S) \simeq 8.86$  for  $n > 70$ ) for both the covering radii and mesh-

<sup>8</sup>Given that  $\mathcal{X}$  is not the unit square, we multiply  $R_\star$  and  $R^\star$  given by (12) and (13) by  $\text{vol}^{1/d}(\mathcal{X}) = \sqrt{3}\pi/2$ .

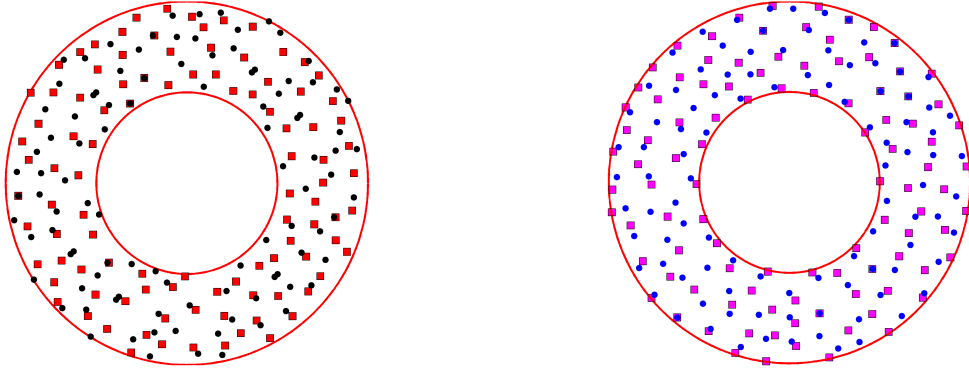


Figure 14: Designs in  $\mathcal{X}$ . Left:  $\mathbf{X}_n^A$  obtained by greedy maximization of  $\widehat{I}_{b_*, B^*, 5}^A(\mathbf{Z}_n)$  (red  $\blacksquare$ ) and points  $\mathbf{X}_n^S$  from Sobol' sequence (black dots). Right: coffee-house designs  $\mathbf{X}_n^{CH, \infty}$  in  $\mathcal{X}$  (magenta  $\blacksquare$ ) and  $\mathbf{X}_n^{CH', \infty}$  in  $\mathcal{X}' \subset \mathcal{X}$  (blue dots).

ratios. The two coffee-house designs  $\mathbf{X}_n^{CH, \infty}$  and  $\mathbf{X}_n^{CH', \infty}$  have larger packing radii than  $\mathbf{X}_n^A$  and thus smaller mesh-ratios (values smaller than one are due to the non-convexity of  $\mathcal{X}$ ).  $\mathbf{X}_n^A$  and  $\mathbf{X}_n^{CH, \infty}$  have smaller covering radii than  $\mathbf{X}_n^{CH', \infty}$  for large values of  $n$ : the absence of sharp corners and edges seems to soften the importance of staying away from the boundaries of the domain; notice that  $\mathbf{X}_n^{CH, \infty}$  has many points near the boundary while  $\mathbf{X}_n^A$  has few, suggesting that boundaries are not necessarily the main concern here.

Overall, comparing the results in Examples 4 and 5, we conclude that the geometry of  $\mathcal{X}$  is clearly a key factor for choosing a suitable  $\beta$  in  $P_\beta(\mathbf{Z}_n)$ , or a suitable erosion of  $\mathcal{X}$  for  $\beta = \infty$  (which should also depend on the maximum design size  $n_2$  envisaged). The choice of  $b_*$  and  $B^*$  in  $\widehat{I}_{b_*, B^*, 5}^A(\mathbf{Z}_n)$  appears to be less critical depending only the volume of  $\mathcal{X}$  (which can always be estimated by Monte Carlo), indicating that greedy optimization of  $\widehat{I}_{b, B, q}^A(\mathbf{Z}_n)$  is a more robust incremental design methodology.

## 6 Discussion

Standard coffee-house design is hard to outperform in terms of mesh-ratio, as, by construction,  $\rho(\mathbf{X}_n) \leq 2$ . However, relying on mesh-ratio alone to measure the space-filling quality of a design is problematic, given its high sensitivity to design packing (or separating) radius (since the mesh-ratio tends to infinity as design points get arbitrarily closer). For that reason, this paper puts a strong emphasis on design methodologies able to guarantee small covering radius, and we believe that the integrated covering measure  $I_{b, B, q}(\mathbf{Z}_n)$  proposed in the paper is an interesting contribution with respect to this goal. As the examples given in the paper show, incremental maximization of this new design criterion is an attractive alternative to classical coffee-house design, leading to design sequences with small covering radii. Substitution of a finite candidate set  $\mathcal{X}_C$  for  $\mathcal{X}$ , accompanied by a lazy-greedy implementation, enables computationally efficient implementations. The major limitation of the proposed methodology is related to the approxi-

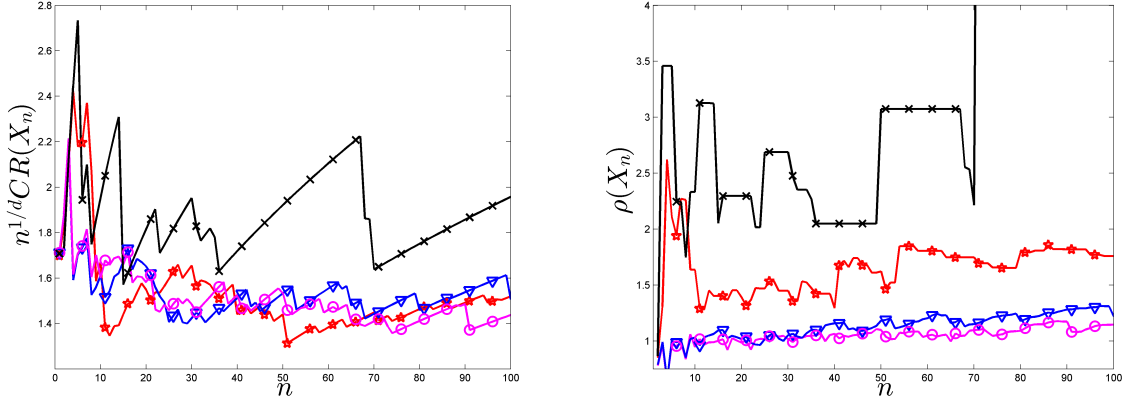


Figure 15:  $n^{1/d} \text{CR}(\mathbf{X}_n)$  (left) and  $\rho_{\mathcal{X}_N}(\mathbf{X}_n)$  (right) for the four designs of Figure 14:  $\mathbf{X}_n^A$  (red  $\star$ ),  $\mathbf{X}_n^S$  (black  $\times$ ),  $\mathbf{X}_n^{CH, \infty}$  (magenta  $\circ$ ) and  $\mathbf{X}_n^{CH', \infty}$  (blue  $\nabla$ ).

mation of the uniform measure on  $\mathcal{X}$  by a discrete one supported on a finite number of points, whose quality is necessarily poor when the dimension  $d$  is very large. Examples with  $\mathcal{X}$  the hypercube  $\mathcal{C}_{10}$  and with a non-convex  $\mathcal{X}$  show that the method can handle generic topological cases, being competitive compared to coffee-house design and its variants that aim at keeping design points away from the boundary.

The paper shows that variants of coffee-house design based on spacings have moderate computational complexity, being able to handle large sets of candidate points, which is particularly important when  $d$  is large. However, these methodologies require a careful setting of the parameter controlling design density near the domain boundary, which must reflect the particular geometric characteristics of each case. On the contrary, the integrated covering criterion  $I_{b,B,q}(\mathbf{Z}_n)$  proposed in the paper automatically handles treatment of the domain boundary. This flexibility comes at the cost of an increased numerical complexity which, nevertheless, remains affordable over a wide range of values of  $d$  and  $n$ .

## Appendix

**Two-point CR-optimal design in  $\mathcal{C}_d$ .** The CR-optimal one-point design is  $\mathbf{Z}_1^* = \mathbf{1}_d/2$ , i.e., the center of  $\mathcal{C}_d$ .

Consider the design  $\mathbf{Z}_2(\alpha) = \{\mathbf{z}_1(\alpha), \mathbf{z}_2(\alpha) = \mathbf{1}_d - \mathbf{z}_1(\alpha)\}$  with  $\mathbf{z}_1(\alpha) = (1/2, \dots, 1/2, \alpha)$ ,  $\alpha < 1/2$ . It defines a partition of  $\mathcal{C}_d$  into two polyhedral Voronoi cells  $\mathcal{C}_1$  and  $\mathcal{C}_2$  which are separated by the bisecting hyperplane  $\mathcal{H}$  of the segment joining  $\mathbf{z}_1(\alpha)$  and  $\mathbf{z}_2(\alpha)$ ,  $\mathcal{H} = \{\mathbf{z} \in \mathcal{C}_d : z_d = 1/2\}$ . We have  $\text{CR}(\mathbf{Z}_2(\alpha)) = \|\mathbf{x}^* - \mathbf{z}_1(\alpha)\|$  for  $\mathbf{x}^*$  a vertex of  $\mathcal{C}_d$  when  $\alpha \geq 1/4$  and for  $\mathbf{x}^*$  having all its coordinates in  $\{0, 1\}$  except the last one equal to  $1/2$  otherwise, implying  $\text{CR}(\mathbf{Z}_2(\alpha)) = \max\{[(d-1)/4 + \alpha^2]^{1/2}, [(d-1)/4 + (1/2 - \alpha)^2]^{1/2}\}$ , whose minimum value is reached for  $\alpha = 1/4$ , with  $\text{CR}(\mathbf{Z}_2(1/4)) = (1/2) \sqrt{d-3}/4 < \sqrt{d}/2$ .

Let  $\mathbf{Z}_2^* = \{\mathbf{z}_1^*, \mathbf{z}_2^*\}$  denote a 2-point CR-optimal design, and denote by  $\mathcal{C}_1^*$  and  $\mathcal{C}_2^*$  the two cells of the Voronoi tessellation of  $\mathcal{C}_d$  that  $\mathbf{Z}_2^*$  generates. Each of them contains exactly  $2^{d-1}$  vertices of

$\mathcal{C}_d$ , since otherwise one of them,  $\mathcal{C}_1^*$  say, would contain two opposite vertices, implying  $\text{CR}(\mathbf{Z}_2^*) \geq \text{diam}(\mathcal{C}_1^*)/2 = \sqrt{d}/2 > \text{CR}(\mathbf{Z}_2(1/4))$ . We have  $\text{CR}(\mathbf{Z}_2^*) \geq \max\{\text{diam}(\mathcal{C}_1^*)/2, \text{diam}(\mathcal{C}_2^*)/2\}$ . The minimum is obtained when the two cells have the same diameter and are such that this diameter is minimal, which gives  $\text{CR}(\mathbf{Z}_2^*) = \text{CR}(\mathbf{Z}_2(1/4))$ .

Consider now an incremental greedy construction, one-step-ahead optimal. Since any 2-point design of the form  $\mathbf{Z}_2 = \{\mathbf{1}_d/2, \mathbf{z}\}$  is such that  $\text{CR}(\mathbf{Z}_2) = \sqrt{d}/2$ , we necessarily have  $\text{CR}(\mathbf{X}_2) = \sqrt{d}/2 > \text{CR}(\mathbf{Z}_2^*)$ . Note that if we could add two points at a time, we could reach  $\text{CR}(\{\mathbf{1}_d/2, \mathbf{x}_2, \mathbf{x}_3\}) = \text{CR}(\mathbf{Z}_2^*)$  by taking  $\mathbf{x}_2 = \mathbf{z}_1^*$ ,  $\mathbf{x}_3 = \mathbf{z}_2^*$ , and even  $\text{CR}(\{\mathbf{1}_d/2, \widehat{\mathbf{x}}_2, \widehat{\mathbf{x}}_3\}) = (1/2)\sqrt{d-8/9}$  by choosing  $\widehat{\mathbf{x}}_2 = (1/2, \dots, 1/2, 1/6)$  and  $\widehat{\mathbf{x}}_3 = \mathbf{1}_d - \widehat{\mathbf{x}}_2$ .

**Evaluation of  $\text{CR}(\mathbf{x}_n)$  when  $\mathcal{X} = \mathcal{C}_d$ .** For  $d \leq 4$ , the exact values of  $\text{CR}(\mathbf{X}_n)$  are calculated using Voronoi tessellations; see [25]; for larger  $d$ , we underestimate  $\text{CR}(\mathbf{X}_n)$  by  $\text{CR}_{\mathcal{X}_N}(\mathbf{X}_n)$  given by (1), with  $\mathcal{X}_N$  a finite subset of  $\mathcal{X}$ . When  $\mathcal{X}_N$  corresponds to the first  $N$  points of a  $(t, s)$ -sequence in base  $\beta$  in  $\mathcal{X}$ , we have

$$\text{CR}(\mathcal{X}_N) < \overline{\text{CR}}(\mathcal{X}_N) = \frac{\sqrt{d} \beta^{1+t/d}}{N^{1/d}}, \quad (23)$$

see [22, Th. 6.11]. The values of  $t$  for the sequences of Sobol' (for which  $\beta = 2$ ) are given in Table 3; see [34]. The sequences of Faure [6] have  $t = 0$  but  $\beta$  is the first prime number larger than or equal to  $d$ , making the upper bound in (23) larger than for Sobol'. Figure 16 shows  $\overline{\text{CR}}(\mathcal{X}_N)$  as a function of  $d$  for both types of sequences when  $N = 10^6$ , showing that the inequality  $\text{CR}(\mathbf{X}_n) \leq \text{CR}_{\mathcal{X}_Q}(\mathbf{X}_n) + \sqrt{d} \beta^{1+t/d} N^{-1/d}$  implied by (2) is of limited practical use. For that reason, for  $d > 4$  we have chosen to only indicate the value of  $\text{CR}_{\mathcal{X}_N}(\mathbf{X}_n)$  for  $\mathcal{X}_N$  given by the union of  $\mathbf{S}_{N_S}$ , the first  $N_S$  points of a (possibly scrambled) Sobol' sequence, and a  $2^d$  full factorial design (i.e., the vertices of  $\mathcal{X}$ ), with thus  $N = N_S + 2^d$ .

Table 3: Values of  $t(d)$  for Sobol' sequence, for  $2 \leq d \leq 13$ .

d	2	3	4	5	6	7	8	9	10	11	12	13
t	0	1	3	5	8	11	15	19	23	27	31	35

**Proof of Theorem 2.** We first prove by induction that  $P_\beta(\mathbf{X}_{n+1}) = (1/2) S_\beta(\mathbf{X}_n)$  for  $n \geq 1$ . Since  $d(\mathbf{x}_2, \partial\mathcal{X}) \leq d(\mathbf{x}_1, \partial\mathcal{X})$ , we have

$$P_\beta(\mathbf{X}_2) = \frac{1}{2} \min\{\beta d(\mathbf{x}_1, \partial\mathcal{X}), \beta d(\mathbf{x}_2, \partial\mathcal{X}), \|\mathbf{x}_1 - \mathbf{x}_2\|\} = \frac{1}{2} \min\{\beta d(\mathbf{x}_2, \partial\mathcal{X}), \|\mathbf{x}_1 - \mathbf{x}_2\|\} = \frac{1}{2} S_\beta(\mathbf{X}_1),$$

which proves the property for  $n = 1$ . Assume that  $P_\beta(\mathbf{X}_n) = (1/2) S_\beta(\mathbf{X}_{n-1})$ . By construction, we have  $P_\beta(\mathbf{X}_{n+1}) \leq (1/2) S_\beta(\mathbf{X}_n)$ ; we show that  $P_\beta(\mathbf{X}_{n+1}) \geq (1/2) S_\beta(\mathbf{X}_n)$ . The greedy construction of  $\mathbf{X}_{n+1}$  implies that  $d(\mathbf{x}_{n+1}, \mathbf{X}_n) \geq S_\beta(\mathbf{X}_n)$  and  $\beta d(\mathbf{x}_{n+1}, \partial\mathcal{X}) \geq S_\beta(\mathbf{X}_n)$ . Also, any  $\mathbf{x}_i \neq \mathbf{x}_j \in \mathbf{X}_n$  satisfy

$$\|\mathbf{x}_i - \mathbf{x}_j\| \geq 2 P_\beta(\mathbf{X}_n) = S_\beta(\mathbf{X}_{n-1}) \geq S_\beta(\mathbf{X}_n),$$

where the first inequality comes from the definition of  $P_\beta$ , the equality comes from the induction hypothesis and the second inequality from the fact that  $S_\beta$  is nonincreasing. Similarly,  $\beta d(\mathbf{x}_i, \partial\mathcal{X}) \geq 2 P_\beta(\mathbf{X}_n) \geq S_\beta(\mathbf{X}_n)$  for all  $\mathbf{x}_i \in \mathbf{X}_n$ . Altogether, this gives  $P_\beta(\mathbf{X}_{n+1}) \geq (1/2) S_\beta(\mathbf{X}_n)$ .

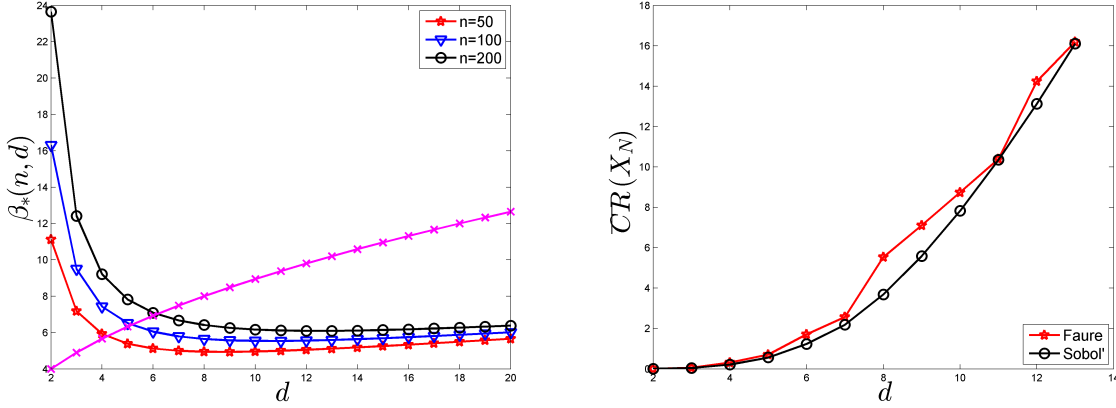


Figure 16: Left:  $\beta_*(n, d)$  given by (21) as a function of  $d$  for  $n = 50, 100, 200$ ; the magenta curve with  $\times$  is for the value  $\beta = 2\sqrt{2d}$  suggested in [33]. Right: upper bound  $CR(\mathcal{X}_N)$  given by (23) as a function of  $d$  for the sequences of Faure and Sobol' for  $N = 10^6$ .

Let  $\mathbf{Z}_n^*$  and  $\mathbf{Z}_n^{**}$  be two  $n$ -point designs in  $\mathcal{X}$  such that  $S_\beta(\mathbf{Z}_n^*) = S_{\beta,n}^* = \min_{\mathbf{Z}_n \subset \mathcal{X}} S_\beta(\mathbf{Z}_n)$  ( $n \geq 1$ ) and  $P_\beta(\mathbf{Z}_n^{**}) = P_{\beta,n}^* = \max_{\mathbf{Z}_n \subset \mathcal{X}} P_\beta(\mathbf{Z}_n)$  ( $n \geq 2$ ). Denote by  $\mathcal{X}^-(r)$  the erosion of  $\mathcal{X}$  by  $\mathcal{B}(\mathbf{0}, r)$ ,  $\mathcal{X}^-(r) = \mathcal{X} \ominus \mathcal{B}(\mathbf{0}, r) = \{\mathbf{x} \in \mathcal{X} : \mathbf{x} + \mathbf{z} \in \mathcal{X}, \forall \mathbf{z} \in \mathcal{B}(\mathbf{0}, r)\}$ , with  $\mathcal{X}^-(r) \neq \emptyset$  for any  $r \leq \text{diam}(\mathcal{X})/2$ .

Suppose that the  $n$  balls  $\mathcal{B}(\mathbf{x}_i, S_\beta(\mathbf{X}_n))$  do not cover  $\mathcal{X}^-(S_\beta(\mathbf{X}_n)/\beta)$ . This would imply the existence of  $\mathbf{x}^* \in \mathcal{X}^-(S_\beta(\mathbf{X}_n)/\beta)$  such that  $\beta d(\mathbf{x}^*, \partial\mathcal{X}) > S_\beta(\mathbf{X}_n)$  and  $d(\mathbf{x}^*, \mathbf{X}_n) > S_\beta(\mathbf{X}_n)$ , and therefore  $D_\beta(\mathbf{x}^*, \mathbf{X}_n) > S_\beta(\mathbf{X}_n)$ . This is impossible since  $S_\beta(\mathbf{X}_n) = \max_{\mathbf{x} \in \mathcal{X}} D_\beta(\mathbf{x}, \mathbf{X}_n)$ .

We prove that  $S_\beta(\mathbf{X}_n) \leq 2S_{\beta,n}^*$ . Take any  $\mathbf{x}_i \in \mathbf{X}_{n+1}$ . It satisfies  $\mathbf{x}_i \in \mathcal{X}^-(2P_\beta(\mathbf{X}_{n+1})/\beta)$  from the definition of  $P_\beta$ , and thus  $\mathbf{x}_i \in \mathcal{X}^-(S_\beta(\mathbf{X}_n)/\beta)$  since  $P_\beta(\mathbf{X}_{n+1}) = (1/2)S_\beta(\mathbf{X}_n)$ , and  $\mathbf{x}_i \in \mathcal{X}^-(S_{\beta,n}^*/\beta)$  since  $S_{\beta,n}^* \leq S_\beta(\mathbf{X}_n)$ . Since the  $n$  balls  $\mathcal{B}(\mathbf{z}_i^*, S_\beta(\mathbf{X}_n))$ ,  $\mathbf{z}_i^* \in \mathbf{Z}_n^*$ , cover  $\mathcal{X}^-(S_{\beta,n}^*/\beta)$ , this implies the existence of a  $\mathbf{z}_\ell^* \in \mathbf{Z}_n^*$  and of  $\mathbf{x}_j \neq \mathbf{x}_k \in \mathbf{X}_{n+1}$  such that  $\mathbf{x}_j, \mathbf{x}_k \in \mathcal{B}(\mathbf{z}_\ell^*, S_{\beta,n}^*)$ . Therefore,  $S_\beta(\mathbf{X}_n) = 2P_\beta(\mathbf{X}_{n+1}) \leq \|\mathbf{x}_i - \mathbf{x}_j\| \leq 2S_{\beta,n}^*$ .

We now prove that  $P_{\beta,n+1}^* \leq 2P_\beta(\mathbf{X}_{n+1})$ . Take any  $\mathbf{z}_i^{**} \in \mathbf{Z}_{n+1}^{**}$ ;  $\mathbf{z}_i^{**} \in \mathcal{X}^-(2P_{\beta,n+1}^*/\beta)$  from the definition of  $P_\beta$ ,  $\mathbf{z}_i^{**} \in \mathcal{X}^-(2P_\beta(\mathbf{X}_{n+1})/\beta) = \mathcal{X}^-(S_\beta(\mathbf{X}_n)/\beta)$  since  $P_{\beta,n+1}^* \geq P_\beta(\mathbf{X}_{n+1}) = (1/2)S_\beta(\mathbf{X}_n)$ , and  $\mathbf{z}_i^{**} \in \mathcal{X}^-(S_{\beta,n}^*/\beta)$  since  $S_{\beta,n}^* \leq S_\beta(\mathbf{X}_n)$ . Therefore, there exist  $\mathbf{z}_\ell^* \in \mathbf{Z}_n^*$  and  $\mathbf{z}_j^{**} \neq \mathbf{z}_k^{**} \in \mathbf{Z}_{n+1}^{**}$  such that  $\mathbf{z}_j^{**}, \mathbf{z}_k^{**} \in \mathcal{B}(\mathbf{z}_\ell^*, S_{\beta,n}^*)$ . It implies that  $2P_{\beta,n+1}^* \leq 2S_{\beta,n}^* \leq 2S_\beta(\mathbf{X}_n) = 4P_\beta(\mathbf{X}_{n+1})$ .

Finally, we show that  $1 \leq \rho_\beta(\mathbf{X}_n) \leq 2$ . Since  $\mathcal{X}$  is convex, for any  $\mathbf{X}_n$ ,  $n \geq 2$ , the  $n$  balls  $\mathcal{B}(\mathbf{x}_i, P_\beta(\mathbf{X}_n))$  do not cover  $\mathcal{X}^-(P_\beta(\mathbf{X}_n)/\beta)$ , which implies the existence of a  $\mathbf{x} \in \mathcal{X}^-(P_\beta(\mathbf{X}_n)/\beta) \setminus \cup_{i=1}^n \mathcal{B}(\mathbf{x}_i, P_\beta(\mathbf{X}_n))$ . It satisfies  $d(\mathbf{x}, \partial\mathcal{X}) \geq P_\beta(\mathbf{X}_n)/\beta$  and  $d(\mathbf{x}, \mathbf{X}_n) \geq P_\beta(\mathbf{X}_n)/\beta$ ; therefore,  $D_\beta(\mathbf{x}, \mathbf{X}_n) \geq P_\beta(\mathbf{X}_n)$  and  $S_\beta(\mathbf{X}_n) \geq P_\beta(\mathbf{X}_n)$ , so that  $\rho_\beta(\mathbf{X}_n) \geq 1$ . Any design obtained with the greedy algorithm satisfies  $S_\beta(\mathbf{X}_{n+1}) \leq S_\beta(\mathbf{X}_n) = 2P_\beta(\mathbf{X}_{n+1})$ ,  $n \geq 1$ , so that  $\rho_\beta(\mathbf{X}_{n+1}) \leq 2$ .  $\blacksquare$



## Acknowledgements

This work was partly supported by project INDEX (INcremental Design of EXperiments) ANR-18-CE91-0007 of the French National Research Agency (ANR).

## Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

## References

- [1] Bach, F.: Learning with submodular functions: a convex optimization perspective. *Foundations and Trends in Machine Learning* **6**(2-3), 145–373 (2013)
- [2] Beck, J., Guillas, S.: Sequential design with mutual information for computer experiments (MICE): emulation of a tsunami model. *SIAM/ASA Journal on Uncertainty Quantification* **4**, 739–766 (2016)
- [3] De Marchi, S.: On optimal center locations for radial basis function interpolation: computational aspects. *Rend. Sem. Math. Torino* **61**(3), 343–358 (2003)
- [4] Fang, K.T., Li, R., Sudjianto, A.: *Design and Modeling for Computer Experiments*. Chapman & Hall/CRC, Boca Raton (2006)
- [5] Cabral Farias, R., Pronzato, L., Rendas, M.J.: Incremental construction of nested designs based on two-level fractional factorial designs. HAL preprint hal-02483004 (2020)
- [6] Faure, H.: Discrepances de suites associées à un système de numération (en dimension  $s$ ). *Acta Arithmetica* **41**, 337–351 (1982)
- [7] Frank, M., Wolfe, P.: An algorithm for quadratic programming. *Naval Res. Logist. Quart.* **3**, 95–110 (1956)
- [8] Gonzalez, T.: Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science* **38**, 293–306 (1985)
- [9] Graf, S., Luschgy, H.: *Foundations of Quantization for Probability Distributions*. Springer, Berlin (2000)
- [10] Janson, S.: Maximal spacings in several dimensions. *The Annals of Probability* **15**(1), 274–280 (1987)
- [11] Johnson, M., Moore, L., Ylvisaker, D.: Minimax and maximin distance designs. *Journal of Statistical Planning and Inference* **26**, 131–148 (1990)
- [12] Kennard, R., Stone, L.: Computer aided design of experiments. *Technometrics* **11**(1), 137–148 (1969)
- [13] Korte, B., Vygen, J.: *Combinatorial Optimization: Theory and Algorithms* (5 ed.). Springer, New York (2012)

- [14] Krause, A., Singh, A., Guestrin, C.: Near-optimal sensor placements in gaussian processes: theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research* **9**, 235–284 (2008)
- [15] Lekivetz, R., Jones, B.: Fast flexible space-filling designs for nonrectangular regions. *Quality and Reliability Engineering International* **31**(5), 829–837 (2015)
- [16] Loepky, J., Sacks, J., Welch, W.: Choosing the sample size of a computer experiment: a practical guide. *Journal of the American Statistical Association* **51**(4), 366–376 (2009)
- [17] Luschgy, H., Pagès, G.: Greedy vector quantization. *Journal of Approximation Theory* **198**, 111–131 (2015)
- [18] Minoux, M.: Accelerated greedy algorithms for maximizing submodular set functions. In: *Proc. 8th IFIP Conference, Wurzburg (Part 2)*, pp. 234–243. Springer, New-York (1977)
- [19] Müller, W.: Coffee-house designs. In: A. Atkinson, B. Bogacka, A. Zhigljavsky (eds.) *Optimum Design 2000*, chap. 21, pp. 241–248. Kluwer, Dordrecht (2001)
- [20] Müller, W.: *Collecting Spatial Data*. Springer, Berlin (2007). [3rd ed.]
- [21] Nemhauser, G., Wolsey, L., Fisher, M.: An analysis of approximations for maximizing submodular set functions–I. *Mathematical Programming* **14**(1), 265–294 (1978)
- [22] Niederreiter, H.: *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia (1992)
- [23] Noonan, J., Zhigljavsky, A.: Efficient quantization and weak covering of high dimensional cubes. arXiv preprint arXiv:2005.07938 (2020)
- [24] Noonan, J., Zhigljavsky, A.: Non-lattice covering and quantization of high dimensional sets. arXiv preprint arXiv:2006.02705v1 (2020)
- [25] 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)
- [26] Pronzato, L., Müller, W.: Design of computer experiments: space filling and beyond. *Statistics and Computing* **22**, 681–701 (2012)
- [27] Pronzato, L., Rendas, J.: A submodular criterion for space-filling design. *SIAM Conference on Uncertainty Quantification, EPFL, Lausanne, Switzerland, April 5-8* (2016)
- [28] Pronzato, L., Zhigljavsky, A.: Measures minimizing regularized dispersion. *J. Scientific Computing* **78**(3), 1550–1570 (2019)
- [29] Pronzato, L., Zhigljavsky, A.: Bayesian quadrature, energy minimization and space-filling design. *SIAM/ASA J. Uncertainty Quantification* **8**(3), 959–1011 (2020)
- [30] 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)

- [31] Santner, T., Williams, B., Notz, W.: The Design and Analysis of Computer Experiments. Springer, Heidelberg (2003)
- [32] Schaback, R.: Error estimates and condition numbers for radial basis function interpolation. *Advances in Computational Mathematics* **3**(3), 251–264 (1995)
- [33] Shang, B., Apley, D.: Full-sequential space-filling design algorithms for computer experiments. *Journal of Quality Technology* (2020). <https://doi.org/10.1080/00224065.2019.1705207>
- [34] Sobol', I.: On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics* **7**(4), 86–112 (1967)
- [35] Wynn, H.: The sequential generation of  $D$ -optimum experimental designs. *Annals of Math. Stat.* **41**, 1655–1664 (1970)
- [36] Zhigljavsky, A., Noonan, J.: Covering of high-dimensional cubes and quantization. arXiv preprint arXiv:2002.06118 (2020)