Available online at www.sciencedirect.com

ScienceDirect

Journal of Approximation Theory 294 (2023) 105931

Journal of Approximation Theory

www.elsevier.com/locate/jat

Full Length Article

Quasi-uniform designs with optimal and near-optimal uniformity constant

L. Pronzato[a], A. Zhigljavsky[b,*]

[a] Université Côte d'Azur, CNRS, Laboratoire I3S, 2000 route des Lucioles, 06900, Sophia Antipolis, France
[b] School of Mathematics, Cardiff University, Senghennydd Road, Cardiff, CF24 4YH, UK

Received 18 December 2021; received in revised form 3 May 2023; accepted 3 June 2023
Available online 14 June 2023

Communicated by R. Schaback

Abstract

A design is a collection of distinct points in a given set $\mathscr{X}$, which is assumed to be a compact subset of $\mathbb{R}^d$, and the mesh-ratio of a design is the ratio of its fill distance to its separation radius. The uniformity constant of a sequence of nested designs is the smallest upper bound for the mesh-ratios of the designs. We derive a lower bound on this uniformity constant and show that a simple greedy construction achieves this lower bound. We then extend this scheme to allow more flexibility in the design construction.

Keywords: Separation radius; Packing radius; Fill distance; Mesh norm; Covering radius; Mesh-ratio; Quasi-uniform design; Greedy algorithm

1. Introduction

Let $\mathscr{X}$ be a compact subset of $\mathbb{R}^d$, for some $d \geq 1$, with $\mathrm{vol}(\mathscr{X}) > 0$.

Let $\|\cdot\|$ denote a norm, not necessarily the Euclidean norm $\|\cdot\|_2$, on $\mathbb{R}^d$. The ball of radius r and center $\mathbf{x}$ is $\mathscr{B}(\mathbf{x}, r) = \{\mathbf{x}' \in \mathbb{R}^d : \|\mathbf{x}' - \mathbf{x}\| \leq r\}$. The volume of the unit ball $\mathscr{B}(\mathbf{0}, 1)$ is denoted by V_d. If the norm $\|\cdot\|$ is Euclidean, then $V_d = \pi^{d/2}/\Gamma(d/2 + 1)$.

A collection $\mathbf{X}_n = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ of n distinct points in $\mathscr{X}$ will be called an n-point design (in the modern literature on approximation theory, designs are often called "data sets", see e.g. [16,20]). We start with several definitions of well-known characteristics of designs.

* Corresponding author.
E-mail addresses: pronzato@i3s.unice.fr (L. Pronzato), ZhigljavskyAA@cardiff.ac.uk (A. Zhigljavsky).

https://doi.org/10.1016/j.jat.2023.105931

FD, *the fill distance* (also known as mesh norm, covering radius, dispersion, or minimax-distance criterion), of the n-point design $\mathbf{X}_n$ for $\mathscr{X}$ is

$$h(\mathbf{X}_n) = h_{\mathscr{X}}(\mathbf{X}_n) = \sup_{\mathbf{x}\in\mathscr{X}} \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}-\mathbf{x}_i\|\,,\ n \geq 1.$$

A design $\mathbf{X}^*_{n,FD}$ will be called FD-optimal if $h^*_n = h(\mathbf{X}^*_{n,FD}) = \min_{\mathbf{X}_n\in\mathscr{X}} h(\mathbf{X}_n)$. SR, *the separation radius* (also called packing radius or maximin-distance criterion), of $\mathbf{X}_n$ is

$$q(\mathbf{X}_n) = \frac{1}{2} \min_{\mathbf{x}_i\neq\mathbf{x}_j\in\mathbf{X}_n} \|\mathbf{x}_i-\mathbf{x}_j\|\,,\ n \geq 2.$$

A design $\mathbf{X}^*_{n,q}$ will be called SR-optimal if $q^*_n = q(\mathbf{X}^*_{n,SR}) = \max_{\mathbf{X}_n\in\mathscr{X}} q(\mathbf{X}_n)$. *The mesh-ratio* of $\mathbf{X}_n$ for $\mathscr{X}$ is

$$\mathsf{MR}(\mathbf{X}_n) = \mathsf{MR}_{\mathscr{X}}(\mathbf{X}_n) = \frac{h_{\mathscr{X}}(\mathbf{X}_n)}{q(\mathbf{X}_n)}\,,\ n \geq 2.$$

The mesh-ratio provides a measure of how uniformly points in $\mathbf{X}_n$ are distributed in $\mathscr{X}$, see e.g. [16, p. 573] and [5, p. 129]; it is sometimes called the uniformity constant of $\mathbf{X}_n$, see [3]. The mesh-ratio is commonly used to assess the stability of approximations constructed on the base of observations at $\mathbf{x}_i \in \mathbf{X}_n$, see e.g. [16] and [20, Chapter 12]. According to Guideline 7.10 in [16, p. 579], the best approximation error with the most stable system is achieved by using quasi-uniform designs (data sets) with the smallest mesh-ratio. The mesh-ratio is fundamental in estimation of stability of approximations through the approach involving the Lebesgue constant, see [3, Th. 1] and [8, Sect. 8.5]. Moreover, the mesh-ratio plays an important role in the derivation of upper bounds on the quality of kernel approximations in the so-called 'escape theorems', when the approximated function is less smooth than the kernel, see [10,11] as well as [5, Th. 1, p. 129] and [16, Th. 7.8].

Let $\mathbf{X}_\infty = \{\mathbf{x}_1, \mathbf{x}_2, \ldots\} \subset \mathscr{X}$ be a sequence of points in $\mathscr{X}$. There is a one-to-one correspondence between such point sequence $\mathbf{X}_\infty$ and the sequence $\{\mathbf{X}_n\}_{n=1}^\infty$ of nested designs $\mathbf{X}_n = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$. A sequence $\{\mathbf{X}_n\}_{n=1}^\infty$ of nested designs $\mathbf{X}_n$ in a compact set $\mathscr{X} \subset \mathbb{R}^d$ is called quasi-uniform if there exists a constant $b < \infty$ such that $\mathsf{MR}(\mathbf{X}_n) \leq b$ for all n. The smallest such $b = \mathsf{MR}(\mathbf{X}_\infty)$ is called the *uniformity constant* of the corresponding sequence of nested designs $\{\mathbf{X}_n\}_{n=1}^\infty$. Quasi-uniform sequences of designs with small uniformity constants are the main sources of designs (point sets) in the meshless (or "mesh-free") methods of computational mathematics; see e.g [5,16,20]. A sequence $\mathbf{X}^*_\infty = \{\mathbf{x}^*_1, \mathbf{x}^*_2, \ldots\}$ will be called MR-optimal if its uniformity constant is minimal:

$$\mathsf{MR}(\mathbf{X}^*_\infty) = \min_{\mathbf{X}_\infty\subset\mathscr{X}} \mathsf{MR}(\mathbf{X}_\infty)\,. \tag{1.1}$$

It is well known that when $\mathscr{X}$ is connected, $\mathsf{MR}(\mathbf{X}_n) \geq 1$ for any n-point design $\mathbf{X}_n$ in $\mathscr{X}$ (as the n-balls $\mathscr{B}(\mathbf{x}_i, \mathsf{CR}(\mathbf{X}_n))$ must cover $\mathscr{X}$). One of the main results of the paper is Theorem 1.1, which states that in fact $\limsup_{n\to\infty} \mathsf{MR}(\mathbf{X}_n) \geq 2$ for any compact $\mathscr{X}$ with positive volume. The proof is rather elementary but the result does not seem to be known. It implies in particular that the classical greedy packing algorithm is MR-optimal.

Theorem 1.1. *For any sequence of nested designs $\mathbf{X}_n$ in a compact set $\mathscr{X} \subset \mathbb{R}^d$ with* $\mathrm{vol}(\mathscr{X}) > 0$, *we have*

$$\limsup_{n\to\infty} \mathsf{MR}(\mathbf{X}_n) \geq 2.$$

In particular, $\mathsf{MR}(\mathbf{X}_\infty) \geq 2$ *for any* $\mathbf{X}_\infty \subset \mathscr{X}$.

Theorem 1.1 is proved in Section 2. The greedy-packing (or coffee-house) algorithm is presented in Section 3.1; it constructs a sequence $\mathbf{X}_\infty$ with $\mathsf{MR}(\mathbf{X}_n) \le 2$ for all $n \ge 2$ and hence $\mathsf{MR}(\mathbf{X}_\infty) = 2$. In Section 3.2, we generalize the greedy-packing algorithm to the construction of other quasi-uniform sequences with bounded $\mathsf{MR}(\mathbf{X}_\infty)$. In Section 3.3 we use the results of Section 3.2 to establish properties of an implementable version of the greedy-packing algorithm where, at every iteration, the next design point $\mathbf{x}_{n+1}$ is chosen among a finite set of candidates $\mathscr{X}_N \subset \mathscr{X}$ rather than within the whole $\mathscr{X}$. In Section 3.4 we consider a boundary-phobic version of greedy packing, which provides designs with worse (larger) mesh-ratio but better (smaller) fill distance. A connection with two greedy kernel-based constructions (energy minimization and the P-greedy algorithm) is presented in Section 3.5. Section 4 briefly concludes. The Matlab scripts used to produce Figs. 1 and 2 are available at https://sdb3.i3s.unice.fr/anrindex/fr/node/5.

2. Proof of Theorem 1.1

Before providing a proof of Theorem 1.1, we prove two simple lemmas, both of them presenting independent interest.

Lemma 2.1. *For any design $\mathbf{X}_n$ in a compact set $\mathscr{X} \subset \mathbb{R}^d$, we have*

$$[\mathrm{vol}(\mathscr{X})/V_d]^{1/d}\, n^{-1/d} \le h(\mathbf{X}_n)\,,\ \ n \ge 1.$$

Moreover, for any m such that $n \ge m \ge 2$, we have

$$q(\mathbf{X}_n) \le [\mathrm{vol}(\mathscr{X}_0)/V_d]^{1/d}\, n^{-1/d},$$

where $\mathscr{X}_0 = \mathscr{X} \oplus \mathscr{B}(\mathbf{0}, q(\mathbf{X}_m))$, $\mathbf{X}_m$ is a sub-design of $\mathbf{X}_n$ consisting of m points and $\oplus$ denotes the Minkowski sum.

Proof. The n balls $\mathscr{B}(\mathbf{x}_i, h(\mathbf{X}_n))$ cover $\mathscr{X}$; this yields the first inequality. The second inequality follows from $q(\mathbf{X}_n) \le q(\mathbf{X}_m)$, which implies that all the balls $\mathscr{B}(\mathbf{x}_i, q(\mathbf{X}_n))$ are fully inside $\mathscr{X}_0$ $(i = 1, \dots, n)$. □

Lemma 2.1 has the following consequence concerning the rate of decrease of the fill distance and separation radius of quasi-uniform sequences of nested designs.

Corollary 2.1. *For any quasi-uniform sequence of nested designs $\mathbf{X}_n$ with uniformity constant b in a compact set $\mathscr{X} \subset \mathbb{R}^d$, we have*

$$c_1\, n^{-1/d} \le h(\mathbf{X}_n) \le b\, q(\mathbf{X}_n) \le c_2\, n^{-1/d}\,,\ \ \forall n \ge 2\,, \tag{2.1}$$

where c_1 and c_2 are some positive constants.

In the case of Euclidean norm, the statement of Corollary 2.1 is proved in [20]; see Proposition 14.1 and the discussion just after it.

Lemma 2.2. *Let, for any given $n \in \mathbb{N}$, $\mathbf{X}_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ and $\mathbf{X}'_{n+1} = \{\mathbf{x}'_1, \dots, \mathbf{x}'_{n+1}\}$ be arbitrary n-point and $(n+1)$-point designs in $\mathscr{X}$. Then*

$$q(\mathbf{X}'_{n+1}) \le h(\mathbf{X}_n).$$

Proof. Since the n balls $\mathscr{B}(\mathbf{x}_i, h(\mathbf{X}_n))$ cover $\mathscr{X}$, the pigeon-hole principle implies that at least one of them must contain at least two points $\mathbf{x}'_i$ and $\mathbf{x}'_j$ from $\mathbf{X}'_{n+1}$. Therefore, $\|\mathbf{x}'_i - \mathbf{x}'_j\| \le 2\,h(\mathbf{X}_n)$, implying $q(\mathbf{X}'_{n+1}) \le h(\mathbf{X}_n)$. □

Proof of Theorem 1.1. Assume that $\limsup_{n\to\infty} \mathsf{MR}(\mathbf{X}_n) < 2$. This would yield that there exist $r < 2$ and n_0 such that $\mathsf{MR}(\mathbf{X}_n) \le r$ for all $n \ge n_0$.

Consider all such $n \ge n_0$. The definition of $h(\mathbf{X}_n)$ and $\mathsf{MR}(\mathbf{X}_n)$ implies the existence of $\mathbf{x}_j \in \mathbf{X}_n$ such that

$$\|\mathbf{x}_{n+1} - \mathbf{x}_j\| \le h(\mathbf{X}_n) \le r\, q(\mathbf{X}_n)\,.$$

Therefore,

$$q(\mathbf{X}_{n+1}) \le (1/2) \min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x}_{n+1} - \mathbf{x}_i\| \le (r/2)\, q(\mathbf{X}_n).$$

This implies the exponential decrease of $q(\mathbf{X}_n)$ to zero (as $n \to \infty$), which contradicts (2.1). □

3. Construction of sequences of quasi-uniform designs

3.1. Greedy packing

Let us first describe the greedy-packing algorithm (called "geometric greedy method" in [4]), which achieves the lower bound of Theorem 1.1 and hence constructs an MR-optimal sequence of points $\mathbf{X}_\infty$ and nested designs $\{\mathbf{X}_n\}_{n=1}^\infty$. This algorithm is sometimes called the "coffee-house" algorithm, due to the analogy with the behavior of customers in large coffee shops, where new clients tend to seat as far as possible from occupied tables [9].

Algorithm 1 (Greedy packing)

Require: $\mathscr{X}$ compact subset of $\mathbb{R}^d$, $\mathbf{x}_1 \in \mathscr{X}$.
1: set $n = 1$, $\mathbf{X}_1 = \{\mathbf{x}_1\}$;
2: for $n = 1, 2, \ldots$ do the following:
3: find $\mathbf{x}_{n+1} \in \operatorname{Arg\,max}_{\mathbf{x}\in\mathscr{X}} \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x} - \mathbf{x}_i\|$,
4: set $\mathbf{X}_{n+1} = \mathbf{X}_n \cup \{\mathbf{x}_{n+1}\}$.

For arbitrary $\mathbf{x}_1 \in \mathscr{X}$ and any choice of $\mathbf{x}_{n+1} \in \operatorname{Arg\,max}_{\mathbf{x}\in\mathscr{X}} \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x} - \mathbf{x}_i\|$ at step 3, the sequence of designs $\mathbf{X}_n$ constructed by Algorithm 1 satisfies the following property.

Lemma 3.1. *For all $n \ge 2$, the designs $\mathbf{X}_n$ generated by Algorithm 1 satisfy $q(\mathbf{X}_n) = h(\mathbf{X}_{n-1})/2$.*

Proof. The inequality $q(\mathbf{X}_n) \ge h(\mathbf{X}_{n-1})/2$ is proved in [4, Lemma 5.1] by induction on n; the equality is obtained by the same arguments.

By the definition of $\mathbf{x}_2$, we have $q(\mathbf{X}_2) = h(\mathbf{X}_1)/2$. Assume that $q(\mathbf{X}_n) = h(\mathbf{X}_{n-1})/2$ and consider $q(\mathbf{X}_{n+1})$:

$$\begin{aligned} q(\mathbf{X}_{n+1}) &= \min\left\{q(\mathbf{X}_n), (1/2) \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}_{n+1} - \mathbf{x}_i\|\right\} \\ &= \min\{q(\mathbf{X}_n), h(\mathbf{X}_n)/2\} \\ &= \min\{h(\mathbf{X}_{n-1})/2, h(\mathbf{X}_n)/2\} = h(\mathbf{X}_n)/2\,. \quad \square \end{aligned}$$

Theorem 3.2. *For all $n \geq 2$, the designs $\mathbf{X}_n$ generated by Algorithm 1 satisfy*

$$h(\mathbf{X}_n) \leq 2\, h^*_n\,, \quad q(\mathbf{X}_n) \geq \frac{1}{2}\, q^*_n\,, \quad \mathsf{MR}(\mathbf{X}_n) \leq 2.$$

Proof. By Lemma 2.2 applied to the designs $\mathbf{X}_{n+1}$ and $\mathbf{X}^*_{n,FD}$, we obtain $q(\mathbf{X}_{n+1}) \leq h^*_n$. Using Lemma 3.1, this gives $h(\mathbf{X}_n) \leq 2\, h^*_n$. From Lemma 2.2 applied to the designs $\mathbf{X}^*_{n+1,SR}$ and $\mathbf{X}_n$ and Lemma 3.1, we obtain $q^*_{n+1} \leq h(\mathbf{X}_n) = 2\, q(\mathbf{X}_{n+1})$. Finally, $\mathsf{MR}(\mathbf{X}_{n+1}) = h(\mathbf{X}_{n+1})/q(\mathbf{X}_{n+1}) \leq h(\mathbf{X}_n)/q(\mathbf{X}_{n+1}) = 2\,.$ □

Theorem 3.2 may be deduced from Theorem 2.2 in [6], where Algorithm 1 is used to minimize the maximum intercluster distance; see also [7, Theorem 4.3]. Theorem 3.2 also follows from Theorem 3.6. However, we think that the proof provided above is interesting in itself, as the important role of Lemma 3.1 uncovers the key property of Algorithm 1.

Note that in Theorem 3.2 the choice of the norm in $\mathscr{X}$ is irrelevant. Moreover, $\mathscr{X}$ does not have to be a subset of $\mathbb{R}^d$; in particular, $\mathscr{X}$ can be a discrete set as in the clustering problems considered in [6].

While the calculation of $q(\mathbf{X}_n)$ is straightforward, $h(\mathbf{X}_n)$ is difficult to compute when $\mathscr{X}$ is a continuous set. Methods of computational geometry can sometimes be used [14], but are restricted to low-dimensional spaces. The substitution of a finite set $\mathscr{X}_N$ for $\mathscr{X}$, with the N points of $\mathscr{X}_N$ suitably well spread over $\mathscr{X}$, is often used in practice; see Section 3.3 for the analysis of this version of Algorithm 1.

For $d = 1$ and $\mathscr{X} = [0, 1]$, Algorithm 1 initialized at $x_1 = 1/2$ is equivalent to the celebrated van der Corput sequence in base 2 in terms of the behavior of $h(\mathbf{X}_n)$, $q(\mathbf{X}_n)$ and $\mathsf{MR}(\mathbf{X}_n)$; see [12, p. 25]. The regular pattern of $\mathsf{MR}(\mathbf{X}_n)$ observed in dimension 1 extends to dimension 2 with $\mathscr{X} = [0, 1]^2$ when $\|\cdot\| = \|\cdot\|_2$ and the algorithm is initialized at the center $(1/2, 1/2)$. This is illustrated on the left panel of Fig. 1: $\mathsf{MR}(\mathbf{X}_n)$ takes two values only, 2 and $\sqrt{2}$. The detailed behavior of the algorithm is as follows.

Theorem 3.3. *For any $n \geq 5$, define $m = m(n) = \lfloor \log_2(\sqrt{n/2 - 1/4} - 1/2) \rfloor$. Then the packing and covering performance of Algorithm 1 with $\|\cdot\| = \|\cdot\|_2$, initialized at the center $(1/2, 1/2)$ of $\mathscr{X} = [0, 1]^2$, is as follows:*

$$\begin{array}{llll}
q(\mathbf{X}_n) = \gamma_m\sqrt{2}/4\,, & h(\mathbf{X}_n) = \gamma_m/2\,, & \mathsf{MR}(\mathbf{X}_n) = \sqrt{2}\,, & \text{for } n = n_m\,,\\
q(\mathbf{X}_n) = \gamma_m/4\,, & h(\mathbf{X}_n) = \gamma_m/2\,, & \mathsf{MR}(\mathbf{X}_n) = 2\,, & \text{for } n = n_m + 1, \ldots, k_m - 1\,,\\
q(\mathbf{X}_n) = \gamma_m/4\,, & h(\mathbf{X}_n) = \gamma_m\sqrt{2}/4\,, & \mathsf{MR}(\mathbf{X}_n) = \sqrt{2}\,, & \text{for } n = k_m\,,\\
q(\mathbf{X}_n) = \gamma_m\sqrt{2}/8\,, & h(\mathbf{X}_n) = \gamma_m\sqrt{2}/4\,, & \mathsf{MR}(\mathbf{X}_n) = 2\,, & \text{for } n = k_m + 1, \ldots, n_{m+1} - 1\,,
\end{array}$$

where $\gamma_m = 2^{-m}$, $n_m = (2^m + 1)^2 + 4^m$ and $k_m = (2^{m+1} + 1)^2$.

For the sake of brevity, we only give a sketch of the full proof. It is based on the self-replicating pattern of the construction. The first five points in $\mathscr{X} = [0, 1]^2$ correspond to the corners and the center of the square. This gives the initialization for the beginning of the initial cycle, indexed by $m = 0$, with m denoting the cycle number. Define the initialization of cycle m as the replication of the initial design of cycle 0 into 4^m squares of side length $\gamma_m = 2^{-m}$, which form a regular partition of $[0, 1]^2$. The initial design for cycle m has thus $n_m = (2^m + 1)^2 + 4^m = 2^{2m+1} + 2^{m+1} + 1$ points: $(2^m + 1)^2$ of them form a regular grid of width γ_m (i.e., a $(2^m + 1)^2$ full factorial design); the other 4^m points are the centers of the small squares. When moving to the next cycle, the algorithm first (*i*) adds the midpoints of the sides of all small squares (in arbitrary order), then (*ii*) adds the 4^{m+1} centers of the smaller squares created at previous phase. The number of points added during phase (*i*) equals

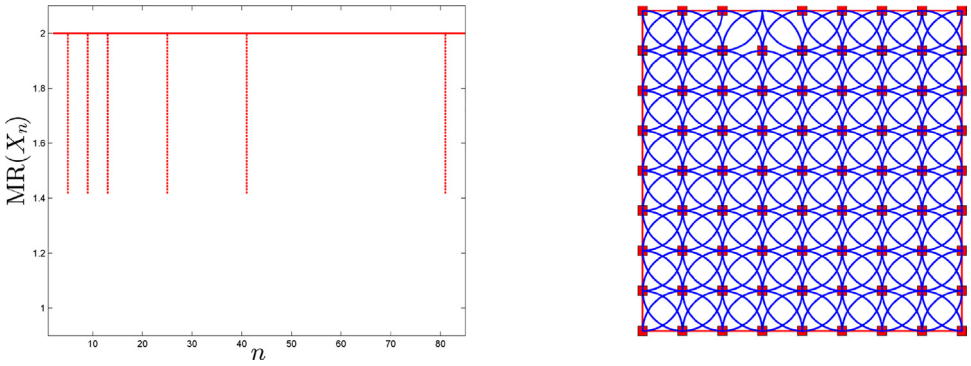


Fig. 1. Designs generated by Algorithm 1 in $\mathscr{X} = [0,1]^2$ with $\|\cdot\| = \|\cdot\|_2$ and $\mathbf{x}_1 = (1/2, 1/2)$. Left: $\mathsf{MR}(\mathbf{X}_n)$ for $n = 2, \ldots, 85$. Right: $\mathbf{X}_{80}$; the circles have radii $h(\mathbf{X}_{80}) = \gamma_2/2 = 0.125$.

$\ell_m = (2^{m+1}+1)^2 - [(2^m+1)^2 + 4^m] = 2^{m+1}(2^m+1)$. For any $n \geq 5$, the associated cycle number $m = m(n)$ is the unique integer satisfying $n_m \leq n < n_{m+1}$. As $n_m = 2(2^m + 1/2)^2 + 1/2$, this gives $m(n) = \lfloor \log_2(\sqrt{n/2 - 1/4} - 1/2) \rfloor$.

Example 3.4. We take $\mathscr{X} = [0,1]^2$, $\|\cdot\| = \|\cdot\|_2$ and $\mathbf{x}_1 = (1/2, 1/2)$. Algorithm 1 progressively imbeds regular grids in $\mathscr{X}$. The left panel of Fig. 1 shows the evolution of $\mathsf{MR}(\mathbf{X}_n)$ as a function of $n = 2, \ldots, 85$; the right panel shows $\mathbf{X}_n$ for $n = 80 = k_2 - 1$.

The regular pattern observed on $[0,1]^d$ for $d = 1, 2$ is maintained for $d = 4$, and Algorithm 1 has the following behavior in $[0,1]^4$.

Theorem 3.5. *For any $n \geq 17$, define $m = m(n)$ as the unique integer satisfying $n_m \leq n < n_{m+1}$, with $n_m = (2^m+1)^4 + 2^{4m}$. Then the packing and covering performance of Algorithm 1 with $\|\cdot\| = \|\cdot\|_2$, initialized at the center $(1/2, 1/2, 1/2, 1/2)$ of $\mathscr{X} = [0,1]^4$, is as follows:*

$$
\begin{aligned}
&q(\mathbf{X}_n) = \gamma_m/2\,, && h(\mathbf{X}_n) = \gamma_m\sqrt{2}/2\,, && \mathsf{MR}(\mathbf{X}_n) = \sqrt{2}\,, && \text{for } n = n_m\,,\\
&q(\mathbf{X}_n) = \gamma_m/(2\sqrt{2})\,, && h(\mathbf{X}_n) = \gamma_m\sqrt{2}/2\,, && \mathsf{MR}(\mathbf{X}_n) = 2\,, && \text{for } n = n_m+1, \ldots, n_m+\ell_m-1\,,\\
&q(\mathbf{X}_n) = \gamma_m/(2\sqrt{2})\,, && h(\mathbf{X}_n) = \gamma_m/2\,, && \mathsf{MR}(\mathbf{X}_n) = \sqrt{2}\,, && \text{for } n = n_m+\ell_m\,,\\
&q(\mathbf{X}_n) = \gamma_m/4\,, && h(\mathbf{X}_n) = \gamma_m/2\,, && \mathsf{MR}(\mathbf{X}_n) = 2\,, && \text{for } n = n_m+\ell_m+1, \ldots, n_{m+1}-1\,,
\end{aligned}
$$

where $\gamma_m = 2^{-m}$ and $\ell_m = 6 \times 2^{2m}\,(2^m+1)^2$.

The proof is omitted. Similarly to the 2-dimensional case treated in Theorem 3.3, the construction follows a self-replicating pattern. The first 17 points in $\mathscr{X} = [0,1]^4$ are the 16 vertices and the center of $\mathscr{X}$. This gives the initialization for the beginning of the cycle $m = 0$, which consists of the following two stages: (*i*) the algorithm chooses (in arbitrary order) all points with two coordinates equal to 1/2 and the other two coordinates in $\{0, 1\}$; there are $2^2 \times \binom{4}{2} = 24$ such points; (*ii*) the algorithm chooses (in arbitrary order) points with one coordinate 1/2 and the other three in $\{0, 1\}$ (there are $2^3 \times \binom{4}{1} = 32$ such points), points with three coordinates 1/2 and one in $\{0, 1\}$ (there are $2 \times \binom{4}{3} = 8$ such points) and points with coordinates in $\{1/4, 3/4\}$ (there are 16 such points).

The initialization of cycle m is defined as the replication of the initial design of cycle 0 into 2^{4m} hypercubes of side length $\gamma_m = 2^{-m}$, which form a regular partition of $[0,1]^4$. The initial

design for cycle m has thus $n_m = (2^m+1)^4 + 2^{4m}$ points: $(2^m+1)^4$ of them form a regular grid of width γ_m; the other 2^{4m} points are the centers of the small hypercubes. We thus have 2^{4m} replications of the initial 17-point initial design, but in smaller hypercubes. In each of them, the selections made by the algorithm are similar to those of the cycle $m = 0$.

From the description above, we can observe that the design $\mathbf{X}_{n_m}$ re-scaled by a factor 2^m gives the integer lattice $\mathbb{Z}_4$ truncated to $(i_1, i_2, i_3, i_4) \in \{0, \dots, 2^m\}^4$. Moreover, when $n = n_m + \ell_m$, the design $\mathbf{X}_n$ re-scaled by 2^{m+1} gives the so-called checkerboard lattice D_4 (the subset of the integer lattice $\mathbb{Z}_4$ consisting of quadruples whose sum is even), truncated to $(i_1, i_2, i_3, i_4) \in \{0, \dots, 2^{m+1}\}^4$; note that D_4 is the densest packing lattice in the 4-dimensional space [2, p. 9].

The regular behavior of Algorithm 1 observed for $d = 1$, and $d = 2$ and 4 where the properly re-scaled design $\mathbf{X}_n$ oscillates between the integer point lattice and the checkerboard lattice, does not hold for other dimensions d

3.2. Relaxed greedy packing

We consider now a generalization of Algorithm 1, where the next point at a given iteration is not necessarily the furthest away from current design points, but is guaranteed to be far enough from them. The bounds obtained in Theorem 3.6 are worse than those in Theorem 3.2; however, it can be shown that the relaxation introduced may improve the covering properties of the design sequence generated; see Section 3.4.

Algorithm 2 (Relaxed greedy packing)

Require: $\mathscr{X}$ compact subset of $\mathbb{R}^d$, $\mathbf{x}_1 \in \mathscr{X}$, $a \in (0, 1]$, $\alpha_1, \alpha_2, \dots \in [a, 1]$;
1: set $n = 1$, $\mathbf{X}_1 = \{\mathbf{x}_1\}$;
2: for $n = 1, 2, \dots$ do the following:
3: take any $\mathbf{x}'$ such that $\min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x}' - \mathbf{x}_i\| \geq \alpha_n\, h(\mathbf{X}_n)$ and set $\mathbf{x}_{n+1} = \mathbf{x}'$;
4: set $\mathbf{X}_{n+1} = \mathbf{X}_n \cup \{\mathbf{x}_{n+1}\}$.

At step 3, the choice of $\mathbf{x}_{n+1}$ is arbitrary provided it satisfies the condition indicated. Due to this flexibility, several existing algorithms form particular cases of Algorithm 2, which in fact defines a whole family of algorithms. In particular, one may first select $\mathbf{x}^* \in \operatorname{Arg\,max}_{\mathbf{x} \in \mathscr{X}} \min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x} - \mathbf{x}_i\|$ and then take any point $\mathbf{x}_{n+1} \in \mathscr{B}(\mathbf{x}^*, (1-\alpha_n) h(\mathbf{X}_n))$. Random designs with guaranteed covering and packing performance can easily be generated in this way: for instance, take $\mathbf{x}_{n+1} = (1-\alpha_n)\mathbf{x}_n^* + \alpha_n \mathbf{x}^*$ with $\mathbf{x}_n^* \in \operatorname{Arg\,min}_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x}^* - \mathbf{x}_i\|$ and α_n a random variable (e.g., uniform) in $[a, 1]$, $a > 0$.

Theorem 3.6. *For all $n \geq 2$, the designs $\mathbf{X}_n$ generated by any version of Algorithm 2 satisfy*

$$h(\mathbf{X}_n) \leq \frac{2}{a}\, h_n^*, \quad q(\mathbf{X}_n) \geq \frac{a}{2}\, q_n^*, \quad \mathsf{MR}(\mathbf{X}_n) \leq \frac{2}{a}.$$

Proof. We first prove by induction that for all $n \geq 2$, $q(\mathbf{X}_n) \geq (a/2)\, h(\mathbf{X}_{n-1})$.

For $n = 2$, by construction we have $q(\mathbf{X}_2) \geq (\alpha_1/2)\, h(\mathbf{X}_1) \geq (a/2)\, h(\mathbf{X}_1)$.

Assume that $q(\mathbf{X}_n) \geq (a/2)\, h(\mathbf{X}_{n-1})$ and consider $q(\mathbf{X}_{n+1})$. The induction assumption gives

$$\begin{aligned} q(\mathbf{X}_{n+1}) &= \min \left\{ q(\mathbf{X}_n), (1/2) \min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x}_{n+1} - \mathbf{x}_i\| \right\} \\ &\geq \min \{ q(\mathbf{X}_n), (\alpha_n/2)\, h(\mathbf{X}_n) \} \end{aligned}$$

$$\geq \min\{(a/2)\,h(\mathbf{X}_{n-1}), (a/2)\,h(\mathbf{X}_n)\} = (a/2)\,h(\mathbf{X}_n)\,.$$

The inequality proved by induction implies

$$\mathsf{MR}(\mathbf{X}_n) = h(\mathbf{X}_n)/q(\mathbf{X}_n) \leq h(\mathbf{X}_{n-1})/q(\mathbf{X}_n) \leq 2/a.$$

Next, by Lemma 2.2, $h(\mathbf{X}_{n-1}) \geq q_n^*$, and therefore $q(\mathbf{X}_n) \geq (a/2)\,h(\mathbf{X}_{n-1}) \geq (a/2)\,q_n^*$. The same lemma implies $h_{n-1}^* \geq q(\mathbf{X}_n) \geq (a/2)\,h(\mathbf{X}_{n-1})$. □

Theorem 3.2 follows from Theorem 3.6 by taking $a = 1$. As in Theorem 3.2, the choice of the norm in $\mathscr{X}$ is irrelevant and $\mathscr{X}$ does not have to be a subset of $\mathbb{R}^d$.

The following property is an extension of Theorem 3.6 to the situation where $\liminf_{n\to\infty} \alpha_n = a \in (0, 1]$ in Algorithm 2 (i.e., not all α_i are bounded from below by a).

Theorem 3.7. *Suppose that in Algorithm 2 the choice of* $\mathbf{x}_{n+1}$ *at Step 3 is such that the scalars*

$$\alpha_n = \min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x}_{n+1} - \mathbf{x}_i\| / h(\mathbf{X}_n)$$

satisfy $\alpha_n \geq \alpha > 0$ *for all* n *and* $\liminf_{n\to\infty} \alpha_n = a \in [\alpha, 1]$. *Then, the designs* $\mathbf{X}_n$ *satisfy*

$$\limsup_{n\to\infty} \frac{h(\mathbf{X}_n)}{h_n^*} \leq \frac{2}{a}\,, \quad \liminf_{n\to\infty} \frac{q(\mathbf{X}_n)}{q_n^*} \geq \frac{a}{2}\,, \quad \limsup_{n\to\infty} \mathsf{MR}(\mathbf{X}_n) \leq \frac{2}{a}.$$

Proof. As $\liminf_{n\to\infty} \alpha_n = a \in (0, 1]$, for all $\epsilon > 0$, there exists n_0 such that $\alpha_n > a - \epsilon$ for all $n \geq n_0$. We first prove that there exists an $n_1 \geq n_0$ such that $q(\mathbf{X}_{n_1}) > [(a-\epsilon)/2]\,h(\mathbf{X}_{n_1})$. Suppose that this is wrong; that is, $q(\mathbf{X}_n) \leq [(a-\epsilon)/2]\,h(\mathbf{X}_n)$ for all $n \geq n_0$. We get

$$\begin{aligned} q(\mathbf{X}_{n+1}) &= \min\left\{q(\mathbf{X}_n),\ (1/2)\min_{\mathbf{x}_i \in \mathbf{X}_n} \|\mathbf{x}_{n+1} - \mathbf{x}_i\|\right\} \\ &= \min\{q(\mathbf{X}_n),\ (\alpha_n/2)\,h(\mathbf{X}_n)\} \\ &\geq \min\{q(\mathbf{X}_n),\ [(a-\epsilon)/2]\,h(\mathbf{X}_n)\} = q(\mathbf{X}_n)\,, \end{aligned}$$

and thus $q(\mathbf{X}_{n+1}) = q(\mathbf{X}_n)$ for all $n \geq n_0$. As $q(\mathbf{X}_{n_0}) \geq (\alpha/2)\,q_{n_0}^* > 0$ from Theorem 3.6, this is in contradiction with Lemma 2.1 which states that $q(\mathbf{X}_n) \leq Cn^{-1/d}$ for some $C > 0$.

Take now $n = n_1$ such that $q(\mathbf{X}_{n_1}) > [(a-\epsilon)/2]\,h(\mathbf{X}_{n_1})$. We have

$$\begin{aligned} q(\mathbf{X}_{n_1+1}) &\geq \min\left\{q(\mathbf{X}_{n_1}),\ [(a-\epsilon)/2]\,h(\mathbf{X}_{n_1})\right\} \\ &= [(a-\epsilon)/2]\,h(\mathbf{X}_{n_1}) \geq [(a-\epsilon)/2]\,h(\mathbf{X}_{n_1+1})\,, \end{aligned}$$

and therefore by induction

$$q(\mathbf{X}_{n+1}) \geq [(a-\epsilon)/2]\,h(\mathbf{X}_n) \geq [(a-\epsilon)/2]\,h(\mathbf{X}_{n+1})$$

for all $n \geq n_1$. Similarly to the proof of Theorem 3.6, $h(\mathbf{X}_{n-1}) \geq q_n^*$ and $h_{n-1}^* \geq q(\mathbf{X}_n)$ respectively imply that $h_n^*/h(\mathbf{X}_n) \geq (a-\epsilon)/2$ and $q(\mathbf{X}_n)/q_n^* \geq (a-\epsilon)/2$, with, moreover, $\mathsf{MR}(\mathbf{X}_n) \leq 2/(a-\epsilon)$, for all $n > n_1$. As ϵ is arbitrary, the result follows. □

In the next section, Theorem 3.6 is used for assessing properties of an easily implementable version of Algorithm 1, where $\mathbf{x}_{n+1}$ at step 3 is chosen from a finite set.

3.3. Greedy packing for a finite candidate set

Consider a version of Algorithm 1 where $\mathbf{x}_{n+1}$ is chosen among a finite set of candidates $\mathscr{X}_N \subset \mathscr{X}$ rather than from the whole $\mathscr{X}$. This assumption makes the implementation of

Algorithm 1 much simpler but naturally deteriorates its performance. Such implementation of Algorithm 1 can be considered as a special case of Algorithm 2, and hence, as we show below in Theorem 3.9, its performance over entire $\mathscr{X}$ can be assessed. Note that the total number of iterations must be smaller than N, the number of candidate points: indeed, for $n \geq N$, the algorithm degenerates as several points necessarily coincide in $\mathbf{X}_{N+j}$, $j \geq 1$.

Lemma 3.8. *For any n-point design $\mathbf{X}_n$ and any N-point set $\mathscr{X}_N \subset \mathscr{X}$ we have*

$$h_{\mathscr{X}_N}(\mathbf{X}_n) \leq h_{\mathscr{X}}(\mathbf{X}_n) \leq h_{\mathscr{X}_N}(\mathbf{X}_n) + h_{\mathscr{X}}(\mathscr{X}_N).$$

Proof. The inequality $h_{\mathscr{X}_N}(\mathbf{X}_n) \leq h_{\mathscr{X}}(\mathbf{X}_n)$ follows from $\mathscr{X}_N \subset \mathscr{X}$. Next, denoting $\mathscr{X}_N = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}\}$, we have

$$\begin{aligned}
h_{\mathscr{X}}(\mathbf{X}_n) &\leq \sup_{\mathbf{x}\in\mathscr{X}} \min_{\mathbf{x}_i\in\mathbf{X}_n} \min_{\mathbf{x}^{(j)}\in\mathscr{X}_N} \left(\|\mathbf{x}-\mathbf{x}^{(j)}\| + \|\mathbf{x}^{(j)}-\mathbf{x}_i\|\right) \\
&= \sup_{\mathbf{x}\in\mathscr{X}} \left[\min_{\mathbf{x}^{(j)}\in\mathscr{X}_N} \left(\|\mathbf{x}-\mathbf{x}^{(j)}\| + \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}^{(j)}-\mathbf{x}_i\|\right)\right] \\
&\leq \sup_{\mathbf{x}\in\mathscr{X}} \left[\min_{\mathbf{x}^{(j)}\in\mathscr{X}_N} \|\mathbf{x}-\mathbf{x}^{(j)}\| + \max_{\mathbf{x}^{(j)}\in\mathscr{X}_N} \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}^{(j)}-\mathbf{x}_i\|\right] \\
&= \sup_{\mathbf{x}\in\mathscr{X}} \left[\min_{\mathbf{x}^{(j)}\in\mathscr{X}_N} \|\mathbf{x}-\mathbf{x}^{(j)}\| + h_{\mathscr{X}_N}(\mathbf{X}_n)\right] \\
&= h_{\mathscr{X}}(\mathscr{X}_N) + h_{\mathscr{X}_N}(\mathbf{X}_n). \quad \square
\end{aligned}$$

Theorem 3.9. *When Algorithm 1 uses a finite set of candidates $\mathscr{X}_N \subset \mathscr{X}$ and $n < N$, its performance satisfies*

$$\begin{array}{lcll}
h_{\mathscr{X}}(\mathbf{X}_n) & \leq & (2/\alpha_n)\, h^*_n\,, & \forall n \geq 1\,, \\
q(\mathbf{X}_n) & \geq & (\alpha_n/2)\, q^*_n\,, & \forall n \geq 2\,, \\
\mathsf{MR}_{\mathscr{X}}(\mathbf{X}_n) & \leq & 2/\alpha_n\,, & \forall n \geq 2\,,
\end{array} \tag{3.1}$$

with $\alpha_n = 1 - h_{\mathscr{X}}(\mathscr{X}_N)/h_{\mathscr{X}}(\mathbf{X}_n)$.

Proof. Denote $\epsilon = h_{\mathscr{X}}(\mathscr{X}_N)$, so that Lemma 3.8 gives $h_{\mathscr{X}_N}(\mathbf{X}_n) \leq h_{\mathscr{X}}(\mathbf{X}_n) \leq h_{\mathscr{X}_N}(\mathbf{X}_n) + \epsilon$. At step 3 of Algorithm 1, we have

$$\min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}_{n+1} - \mathbf{x}_i\| = h_{\mathscr{X}_N}(\mathbf{X}_n) \geq h_{\mathscr{X}}(\mathbf{X}_n) - \epsilon = \alpha_n\, h_{\mathscr{X}}(\mathbf{X}_n),$$

with $\alpha_n = 1 - \epsilon/h_{\mathscr{X}}(\mathbf{X}_n)$. Since $h_{\mathscr{X}}(\mathbf{X}_n)$ is non-increasing with n, α_n is non-increasing too (it reaches zero when $\mathbf{X}_n$ has exhausted $\mathscr{X}_N$, that is, when $k = N$). Theorem 3.6 with α_n substituted for a implies (3.1). $\square$

As we do not know $h_{\mathscr{X}}(\mathbf{X}_n)$ and thus α_n, we can use the inequality $h_{\mathscr{X}}(\mathbf{X}_n) \geq h_{\mathscr{X}_N}(\mathbf{X}_n)$, which gives $\alpha_n \geq a_n = 1 - h_{\mathscr{X}}(\mathscr{X}_N)/h_{\mathscr{X}_N}(\mathbf{X}_n)$. The inequalities (3.1) then remain true with a_n substituted for α_n, as long as $a_n > 0$.

A result similar to Theorem 3.9 holds when the performance of Algorithm 1 is evaluated on a finite set $\mathscr{X}_{N'} \supset \mathscr{X}_N$ instead of $\mathscr{X}$: we simply substitute $\mathscr{X}_{N'}$ for $\mathscr{X}$ and $\alpha_n = 1 - h_{\mathscr{X}'_N}(\mathscr{X}_N)/h_{\mathscr{X}'_N}(\mathbf{X}_n)$ is evaluated easily.

3.4. Boundary-phobic greedy packing

Versions of the greedy packing algorithm that enforce boundary avoidance have been proposed in [13,17]. There, at iteration $n \geq 2$, the next point $\mathbf{x}_{n+1}$ is chosen in $\operatorname{Arg\,max}_{\mathbf{x}\in\mathscr{X}} D_\beta(\mathbf{x}, \mathbf{X}_n, \mathscr{X})$, where

$$D_\beta(\mathbf{x}, \mathbf{X}_n, \mathscr{X}) = \min\left\{ \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}-\mathbf{x}_i\|,\ \beta\, d(\mathbf{x}, \partial\mathscr{X}) \right\},\ \beta \in (0,\infty), \tag{3.2}$$

with $d(\mathbf{x}, \partial\mathscr{X})$ the distance from $\mathscr{X}$ to the boundary of $\mathscr{X}$. Note that this quantity is easily determined if $\mathscr{X}$ has a simple shape, like a hypercube or a ball, but may be difficult to evaluate otherwise. For $\beta = \infty$, we define $D_\infty(\mathbf{x}, \mathbf{X}_n, \mathscr{X}) = \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}-\mathbf{x}_i\|$ by continuity; the algorithm then coincides with Algorithm 1. For $\beta = 1$, $\mathbf{x}_{n+1}$ is the center of (one of) the largest ball included in $\mathscr{X}$ and not intersecting $\mathbf{X}_n$. For $\beta = 2$, the algorithm corresponds to a greedy method for the solution of the traditional packing problem, for which the n balls do not intersect and are constrained to be fully inside $\mathscr{X}$. For $\beta > 2$, the larger β is, the more the balls are allowed to overshoot $\mathscr{X}$, with their centers remaining inside $\mathscr{X}$. When $\mathscr{X} = [0,1]^d$ and $\|\cdot\| = \|\cdot\|_2$, the value $\beta = 2\sqrt{2d}$ is recommended in [17], while [13] recommends to let β depend on the targeted number $n_{\max}$ of design points and suggests taking

$$\beta = \beta(n_{\max}, d) = \frac{d}{2\,(n_{\max} V_d)^{-1/d}} - \sqrt{d},$$

with $V_d = \pi^{d/2}/\Gamma(d/2+1)$. Both references illustrate the interest of using $\beta < \infty$ instead of Algorithm 1 (where $\beta = \infty$) in terms of fill distance $h(\mathbf{X}_n)$. As shown below, for $\mathscr{X} = [0,1]^d$ the boundary-phobic version of greedy packing becomes a particular case of Algorithm 2.

Theorem 3.10. *For $\mathscr{X} = [0,1]^d$ and $\|\cdot\| = \|\cdot\|_2$, the boundary-phobic algorithm that chooses $\mathbf{x}_{n+1}$ in $\operatorname{Arg\,max}_{\mathbf{x}\in\mathscr{X}} D_\beta(\mathbf{x}, \mathbf{X}_n, \mathscr{X})$ at iteration n, with $D_\beta(\mathbf{x}, \mathbf{X}_n, \mathscr{X})$ defined by (3.2) and $\beta \in (0,\infty)$, forms a particular instance of Algorithm 2 with $\alpha_n = a = 1/(1+\sqrt{d}/\beta)$.*

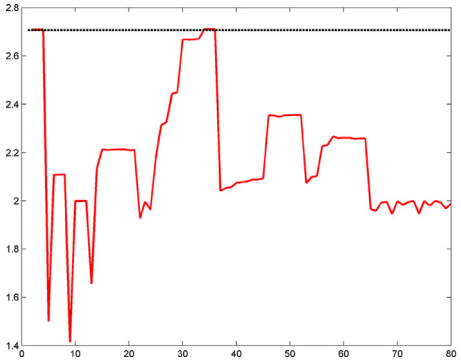
Proof. Let $\mathscr{X} = [0,1]^d$ and $\beta \in (0,\infty)$, $r_n = D_\beta(\mathbf{x}_{n+1}, \mathbf{X}_n, \mathscr{X}) = \max_{\mathbf{x}\in\mathscr{X}} D_\beta(\mathbf{x}, \mathbf{X}_n, \mathscr{X})$. Any $\mathbf{x} \in \mathscr{X}$ satisfies at least one of the two inequalities

$$\min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}-\mathbf{x}_i\| \leq r_n\ ,\ \ d(\mathbf{x}, \partial\mathscr{X}) \leq r_n/\beta.$$

This implies that $\mathscr{X} \setminus \{\mathbf{x} \in \mathbb{R}^d : d(\mathbf{x}, \partial\mathscr{X}) \leq r_n/\beta\} \subset \cup_{i=1}^k \mathscr{B}(\mathbf{x}_i, r_n)$. The inequalities $r_n \leq \beta\, d(\mathbf{x}_{n+1}, \partial\mathscr{X}) \leq \beta/2$ imply that $2\,r_n/\beta \leq 1$, and the set $\mathscr{X} \setminus \{\mathbf{x} \in \mathbb{R}^d : d(\mathbf{x}, \partial\mathscr{X}) \leq r_n/\beta\}$ is a hypercube $\mathcal{C}_n$ with side length $1-2\,r_n/\beta$. This hypercube is covered by the n balls $\mathscr{B}(\mathbf{x}_i, r_n)$, implying that

$$\begin{aligned} h_{\mathscr{X}}(\mathbf{X}_n) &= \sup_{\mathbf{x}\in\mathscr{X}} \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}-\mathbf{x}_i\| \\ &\leq \sup_{\mathbf{x}\in\mathscr{X}} \left[\inf_{\mathbf{x}'\in\mathcal{C}_n} \left(\|\mathbf{x}-\mathbf{x}'\| + \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}'-\mathbf{x}_i\| \right) \right] \\ &\leq \sup_{\mathbf{x}\in\mathscr{X}} \inf_{\mathbf{x}'\in\mathcal{C}_n} \|\mathbf{x}-\mathbf{x}'\| + r_n \leq \sqrt{d}\,(r_n/\beta) + r_n\,. \end{aligned}$$

Since, by definition, $r_n \leq \min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}_{n+1}-\mathbf{x}_i\|$, we have



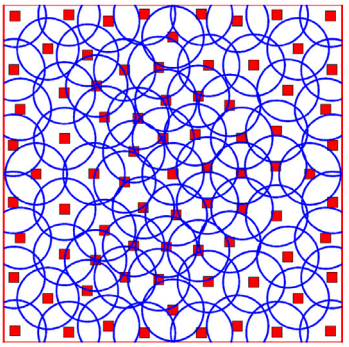


Fig. 2. Designs generated by $\mathbf{x}_{n+1} \in \operatorname{Arg\,max}_{\mathbf{x}\in\mathscr{X}} D_4(\mathbf{x}, \mathbf{X}_n, \mathscr{X})$ in $\mathscr{X} = [0,1]^2$ with $\|\cdot\| = \|\cdot\|_2$ and $\mathbf{x}_1 = (1/2, 1/2)$. Left: $\mathsf{MR}(\mathbf{X}_n)$ for $n = 2, \ldots, 80$; the horizontal line indicates the upper bound $2(1 + \sqrt{d}/\beta)$ on $\mathsf{MR}(\mathbf{X}_n)$. Right: $\mathbf{X}_{80}$; the circles have radii $h(\mathbf{X}_n) = 0.0913$.

$$\min_{\mathbf{x}_i\in\mathbf{X}_n} \|\mathbf{x}_{n+1} - \mathbf{x}_i\| \geq \frac{h_{\mathscr{X}}(\mathbf{X}_n)}{1 + \sqrt{d}/\beta},$$

and the algorithm is a particular instance of Algorithm 2 with $\alpha_n = a = 1/(1 + \sqrt{d}/\beta)$. □

Theorem 3.10 implies that the performance of this algorithm satisfies the bounds indicated in Theorem 3.6.

Example 3.11. We take $\mathscr{X} = [0,1]^2$, $\|\cdot\| = \|\cdot\|_2$ and $\beta = 4$. The left panel of Fig. 2 shows the evolution of $\mathsf{MR}(\mathbf{X}_n)$ as a function of $n = 2, \ldots, 80$ when $\mathbf{X}_n$ is generated by $\mathbf{x}_{n+1} \in \operatorname{Arg\,max}_{\mathbf{x}\in\mathscr{X}} D_\beta(\mathbf{x}, \mathbf{X}_n, \mathscr{X})$ with $\mathbf{x}_1 = (1/2, 1/2)$; the upper bound $2(1 + \sqrt{d}/\beta)$ on $\mathsf{MR}(\mathbf{X}_n)$ is indicated by a horizontal line. The right panel presents $\mathbf{X}_{80}$: comparison with the right panel of Fig. 1 shows that boundary avoidance has significantly reduced $h(\mathbf{X}_n)$. This reduction is obtained at the detriment of $\mathsf{MR}(\mathbf{X}_n)$ for some $\mathbf{X}_n$, as illustrated by the left panels of the two figures (note, however, that $\mathsf{MR}(\mathbf{X}_{80}) < 2$ in Fig. 2).

3.5. Connection with kernel-based methods

3.5.1. Energy minimization

Consider the Riesz kernel $K_s(\mathbf{x}, \mathbf{x}') = 1/\|\mathbf{x} - \mathbf{x}'\|^s$, $s > 0$, and the discrete energy

$$\mathscr{E}_{K_s}(\mathbf{X}_n) = \frac{2}{n(n-1)} \sum_{1\leq i<j\leq n} K_s(\mathbf{x}_i, \mathbf{x}_j)$$

associated with an n-point design $\mathbf{X}_n = \{\mathbf{X}_1, \ldots, \mathbf{x}_n\}$. A design $\mathbf{X}_n^* \subset \mathscr{X}$ that minimizes $\mathscr{E}_{K_s}(\mathbf{X}_n)$ is called a set of s-Fekete points (the original denomination is for $\mathscr{X}$ being the sphere $\mathcal{S}_2$); see, e.g., [1, Chap. 2] for a thorough exposition of the discrete energy problem and [1, Chap. 4] for the connection with the continuous energy problem. In particular, [1, Prop. 2.1.1] shows that $\mathscr{E}_{K_s}(\mathbf{X}_n^*)$ is non decreasing in n; its limit is called the transfinite diameter of $\mathscr{X}$ and coincide with the Wiener constant, i.e., the minimum value of the continuous energy on $\mathscr{X}$ [1, Th. 4.2.2]. Next theorem shows the strong connection that exists between greedy energy minimization for the Riesz kernel and greedy packing.

Theorem 3.12. *The designs $\mathbf{X}_n$ obtained by greedy minimization of the energy $\mathscr{E}_{K_s}$, i.e., such that $\mathbf{x}_{k+1} \in \operatorname{Arg\,min}_{\mathbf{x}\in\mathscr{X}} \mathscr{E}_{K_s}(\mathbf{X}_k \cup \{\mathbf{x}\})$ for all $k \geq 1$, satisfy*

$$h(\mathbf{X}_n) \leq \frac{2}{n^{-1/s}}\, h_n^*\,, \quad q(\mathbf{X}_n) \geq \frac{n^{-1/s}}{2}\, q_n^*\,, \quad \mathsf{MR}(\mathbf{X}_n) \leq \frac{2}{n^{-1/s}}\,, \ n \geq 2\,. \tag{3.3}$$

By letting $s = s_n$ vary at each iteration in K_s in such a way that $s_n/\log n \to \infty$ as $n \to \infty$, we get

$$\limsup_{n\to\infty} \frac{h(\mathbf{X}_n)}{h_n^*} \leq 2\,, \quad \liminf_{n\to\infty} \frac{q(\mathbf{X}_n)}{q_n^*} \geq \frac{1}{2}\,, \quad \limsup_{n\to\infty} \mathsf{MR}(\mathbf{X}_n) \leq 2\,. \tag{3.4}$$

Proof. Direct calculation gives $\mathbf{x}_{k+1} \in \operatorname{Arg\,min}_{\mathbf{x}\in\mathscr{X}} (1/n) \sum_{i=1}^n K_s(\mathbf{x}, \mathbf{x}_i)$. Since, for $\mathbf{x} \notin \mathbf{X}_n$,

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{\|\mathbf{x} - \mathbf{x}_i\|^s} \leq \max_i \frac{1}{\|\mathbf{x} - \mathbf{x}_i\|^s} \leq \sum_{i=1}^n \frac{1}{\|\mathbf{x} - \mathbf{x}_i\|^s}\,,$$

we have

$$\left[\sum_{i=1}^n \frac{1}{\|\mathbf{x} - \mathbf{x}_i\|^s}\right]^{-1/s} \leq \left[\max_i \frac{1}{\|\mathbf{x} - \mathbf{x}_i\|^s}\right]^{-1/s} = \min_i \|\mathbf{x} - \mathbf{x}_i\| \leq n^{1/s} \left[\sum_{i=1}^n \frac{1}{\|\mathbf{x} - \mathbf{x}_i\|^s}\right]^{-1/s}$$

and therefore, for $\mathbf{x}^* \in \operatorname{Arg\,max}_{\mathbf{x}\in\mathscr{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$,

$$\begin{aligned} \min_i \|\mathbf{x}_{n+1} - \mathbf{x}_i\| \geq \left[\sum_{i=1}^n \frac{1}{\|\mathbf{x}_{n+1} - \mathbf{x}_i\|^s}\right]^{-1/s} &\geq \left[\sum_{i=1}^n \frac{1}{\|\mathbf{x}^* - \mathbf{x}_i\|^s}\right]^{-1/s} \\ &\geq n^{-1/s} \min_i \|\mathbf{x}^* - \mathbf{x}_i\| = h(\mathbf{X}_n)\,. \end{aligned}$$

Greedy energy minimization (with the Riesz kernel K_s) thus corresponds to a particular version of relaxed greedy packing (Algorithm 2), where $\alpha_n = n^{-1/s}$ decreases with n, and Theorem 3.6 implies (3.3).

The covering and packing efficiencies of $\mathbf{X}_n$ may degrade as n increases, but if we let $s = s_n$ vary at each iteration in K_s in such a way that $n^{1/s_n} \to 1$ (or equivalently, $s_n/\log n \to \infty$) as $n \to \infty$, then Theorem 3.7 implies (3.4). □

Similar developments with the isotropic Matérn 1/2 kernel with correlation length ℓ, $K_{1/2,\ell}(\mathbf{x}, \mathbf{x}') = \exp(-\|\mathbf{x} - \mathbf{x}'\|/\ell)$, show that $\min_i \|\mathbf{x}_{n+1} - \mathbf{x}_i\| \geq h(\mathbf{X}_n) - \ell \log n$ for greedy energy minimization. Since $h(\mathbf{X}_n) \geq C n^{-1/d}$ for some $C > 0$ from Lemma 2.1, by letting $\ell = \ell_n$ decrease with n in such a way that $n^{1/d}\ell_n \log n \to 0$ as $n \to \infty$, we also get (3.4) from Theorem 3.7.

An advantage of this type of construction over Algorithm 1 is that the choice of $\mathbf{x}_{n+1}$ accounts for the location of all previous $\mathbf{x}_i$, $i \leq n$, and hence is generally uniquely defined, whereas there are often several equivalent choices at step 3 of Algorithm 1 (see for example the regular patterns explained in Theorems 3.3 and 3.5).

3.5.2. Greedy maximum-entropy sampling and the P-greedy algorithm

Consider the kriging framework [19], where an unknown function f on $\mathscr{X}$ is considered as a realization of a Gaussian random field $Z_{\mathbf{x}}$ with zero mean and covariance $\mathsf{E}\{Z_{\mathbf{x}} Z_{\mathbf{x}'}\} = K(\mathbf{x}, \mathbf{x}')$, with K a strictly positive definite kernel defining a reproducing kernel Hilbert space $\mathcal{H}_K$. Let f be evaluated at the n-point design $\mathbf{X}_n = \{\mathbf{X}_1, \ldots, \mathbf{x}_n\}$ and consider a linear predictor $\eta_n(\mathbf{x}_0) = {\mathbf{w}_n}^\top \mathbf{y}_n$ of $f(\mathbf{x}_0)$, with $\mathbf{y}_n = [f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)]^\top$ and $\mathbf{w}_n \in \mathbb{R}^n$. Its Mean-Squared

Error (MSE) is $\mathsf{E}\{|Z_{\mathbf{x}_0} - \mathbf{w}_n{}^\top \mathbf{y}_n|^2\} = \rho_n^2(\mathbf{x}_0, \mathbf{w}_n) = K(\mathbf{x}_0, \mathbf{x}_0) - 2\,\mathbf{w}_n{}^\top \mathbf{k}_n(\mathbf{x}_0) + \mathbf{w}_n{}^\top \mathbf{K}_n \mathbf{w}_n$, where the ith component of the vector $\mathbf{k}_n(\mathbf{x}_0)$ equals $K(\mathbf{x}_0, \mathbf{x}_i)$ and the matrix $\mathbf{K}_n$ has i, j element $K(\mathbf{x}_i, \mathbf{x}_j)$, $i, j = 1, \ldots, n$. The best linear predictor is $\eta_n^*(\mathbf{x}_0) = \mathbf{k}_n{}^\top(\mathbf{x}_0)\mathbf{K}_n^{-1}\mathbf{y}_n$, obtained for $\mathbf{w}_n = \mathbf{K}_n^{-1}\mathbf{k}_n(\mathbf{x}_0)$; its MSE equals $\rho_n^2(\mathbf{x}_0) = K(\mathbf{x}_0, \mathbf{x}_0) - \mathbf{k}_n{}^\top(\mathbf{x}_0)\mathbf{K}_n^{-1}\mathbf{k}_n(\mathbf{x}_0)$. Maximum-entropy sampling [18] constructs an n-point design $\mathbf{X}_n$ by maximizing $\det \mathbf{K}_n$. Since $\det \mathbf{K}_{n+1} = \rho_n^2(\mathbf{x}_{n+1}) \det \mathbf{K}_n$, a greedy version chooses $\mathbf{x}_1 \in \operatorname{Arg\,max}_{\mathbf{x}\in\mathscr{X}} K(\mathbf{x}, \mathbf{x})$ and then $\mathbf{x}_{n+1} \in \operatorname{Arg\,max}_{\mathbf{x}\in\mathscr{X}} \rho_n^2(\mathbf{x})$, $n \geq 1$.

The interpolation of a function $f \in \mathcal{H}_K$ involves the same quantities η_n^* and ρ_n: $\eta_n^*(\cdot)$ is the orthogonal projection of f onto $\operatorname{span}\{K(\cdot, \mathbf{x}_i),\ i = 1, \ldots, n\}$ and $P_n(\cdot) = \rho_n(\cdot)$ is called the power function, with

$$P_n(\mathbf{x}) = \sup_{\|f\|_{\mathcal{H}_K}=1} |f(\mathbf{x}) - \eta_n^*(\mathbf{x})|.$$

In this context, the algorithm for greedy maximum-entropy sampling is called the P-greedy algorithm [15].

When $d = 1$, $K(x, x') = \exp(-|x - x'|/\ell)$ and $\mathscr{X}$ is a closed interval $[a, b]$, the Markov property of the Ornstein–Uhlenbeck process implies that the matrix $\mathbf{K}_n$ is tridiagonal, and for any design point $x_i \in \mathbf{X}_n$ and $x, x' \in [a, b]$ such that $x < x_i < x'$, the random variables Z_x and $Z_{x'}$ are conditionally independent: $\mathsf{E}\{Z_x Z_{x'} | (x_1, Z_{x_1}), \ldots, (x_1, Z_{x_1})\} = K(x, x') - \mathbf{k}_n{}^\top(x)\mathbf{K}_n^{-1}\mathbf{k}_n(x') = 0$. Denote by z_i, $i = 1, \ldots, n$, the reordered design points at iteration n, so that $a \leq z_1 < z_2 < \cdots < z_n \leq b$. Let $\delta_j = z_{j+1} - z_j$, $i = 1, \ldots, n-1$, and $J_n = \operatorname{Arg\,max}_{j=1,\ldots,n-1} \delta_j$ be the set of indices of the most distant neighboring pairs; denote $\delta_n^* = \delta_j$, $j \in J_n$. As $K(x, x) = 1$ for all x, the choice of x_1 is arbitrary. When $x_1 = c = (a + b)/2$, the P-greedy algorithm chooses $x_2 = a$ and $x_3 = b$ (or $x_2 = b$ and $x_3 = a$), and then, for every $n \geq 4$, $\mathbf{x}_{n+1} = (z_j + z_{j+1})/2$ for an arbitrary $j \in J_n$. It thus coincides with the greedy packing algorithm. When $x_1 < c$, the P-greedy algorithm chooses $x_2 = b$ and then, for each n, either (i) $x_{n+1} = (z_j + z_{j+1})/2$ for an arbitrary $j \in J_n$, or (ii) $x_{n+1} = a$. Direct calculation shows that case (ii) occurs when $x_1 - a \geq \Delta_n^* = \delta_n^*/2 - (\ell/2)\log 2 + (\ell/2)\log[1 + \exp(-\delta_n^*/\ell)]$ (with $\Delta_n^* < \delta_n^*/2$), which necessarily holds after a finite number of iterations. For all subsequent iterations, the P-greedy algorithm coincides with greedy packing. The case $x_1 > c$ is treated in a similar way.

The connection with greedy packing is more subtle in other situations ($d > 1$ and/or for other kernels K). Take K isotropic with a small correlation length ℓ. Although it is intuitively clear that the behavior of the P-greedy algorithm will mimic that of greedy packing as long as ℓ is negligible relative to $q(\mathbf{X}_n)$ (so that $\mathbf{K}_n$ is close to the identity matrix), a precise analysis appears difficult. Theorem 15 of [21] indicates that, when K is translation invariant and has finite smoothness τ, the designs $\mathbf{X}_n$ generated by the P-greedy algorithm satisfy $h(\mathbf{X}_n) < cn^{-1/d}$ for some $c > 0$ (note that the rate $n^{-1/d}$ is optimal from Lemma 2.1). The same theorem shows that any relaxation of the P-greedy algorithm that selects an arbitrary $\mathbf{x}_{n+1}$ in the set $\{\mathbf{x} \in \mathscr{X} : P_n(\mathbf{x}) \geq \gamma \max_{\mathbf{x}\in\mathscr{X}} P_n(\mathbf{x})\}$ ($\gamma \in (0, 1]$) also achieves the rate of decrease $n^{-1/d}$ for $h(\mathbf{X}_n)$, and Theorem 19 in the same paper shows that $q(\mathbf{X}_n) > c' n^{-1/d}$ for some $c' > 0$ (so that the rate of decrease of $q(\mathbf{X}_n)$ is optimal too) when $\tau > d/2 + 1$ and $\mathscr{X}$ satisfies an interior cone condition and has a Lipschitz boundary. As for energy minimization (Section 3.5.1), an advantage of this type of construction over Algorithm 1 is that the choice of $\mathbf{x}_{n+1}$ accounts for the location of all previous $\mathbf{x}_i$.

4. Conclusion

By introducing a relaxation in the classical greedy-packing algorithm, we have proposed a general class of simple greedy algorithms that generate quasi-uniform sequences of nested designs with guaranteed packing and covering performance. We have shown that the value 2 of the uniformity constant of the greedy-packing algorithm is optimal, and that it can be attained by a relaxed algorithm whose relaxation vanishes asymptotically. A connection with two kernel-based greedy constructions has been evidenced.

Data availability

No data was used for the research described in the article.

Acknowledgments

The work of the first author was partially funded by project ANR INDEX (ANR-18-CE91-0007).[1] The authors thank an unknown referee for his suggestion to mention the link between greedy packing and the P-greedy algorithm (Section 3.5.2).

References

[1] S.V. Borodachov, D.P. Hardin, E.B. Saff, Discrete Energy on Rectifiable Sets, Springer, 2019.
[2] J.H. Conway, N.J.A. Sloane, Sphere Packings, Lattices and Groups, third ed., Springer, New York, 1999.
[3] S. De Marchi, R. Schaback, Stability of kernel-based interpolation, Adv. Comput. Math. 32 (2) (2010) 155–161.
[4] S. De Marchi, R. Schaback, H. Wendland, Near-optimal data-independent point locations for radial basis function interpolation, Adv. Comput. Math. 23 (3) (2005) 317–330.
[5] G.E. Fasshauer, Meshfree Approximation Methods with MATLAB, World Scientific, 2007.
[6] T.F. Gonzalez, Clustering to minimize the maximum intercluster distance, Theoret. Comput. Sci. 38 (1985) 293–306.
[7] S. Har-Peled, Geometric Approximation Algorithms, American Mathematical Soc., 2011.
[8] A. Iske, Approximation Theory and Algorithms for Data Analysis, Springer, Berlin, 2018.
[9] W.G. Müller, Collecting Spatial Data, third ed., Springer, Berlin, 2007.
[10] F.J. Narcowich, J.D. Ward, H. Wendland, Sobolev bounds on functions with scattered zeros, with applications to radial basis function surface fitting, Math. Comp. 74 (250) (2005) 743–763.
[11] F.J. Narcowich, J.D. Ward, H. Wendland, Sobolev error estimates and a Bernstein inequality for scattered data interpolation with radial basis functions, Constr. Approx. 24 (2006) 175–186.
[12] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, SIAM, Philadelphia, 1992.
[13] A. Nogales Gómez, L. Pronzato, M.-J. Rendas, Incremental space-filling design based on coverings and spacings: improving upon low discrepancy sequences, J. Stat. Theory Pract. 15 (2021) http://dx.doi.org/10.1007/s42519-021-00210-2, in press, article nb. 77.
[14] L. Pronzato, Minimax and maximin space-filling designs: some properties and methods for construction, J. Soc. Fr. Stat. 158 (1) (2017) 7–36.
[15] G. Santin, B. Haasdonk, Convergence rate of the data-independent P-greedy algorithm in kernel-based approximation, Dolomites Res. Notes Approx. 10 (2017) 68–78.
[16] R. Schaback, H. Wendland, Kernel techniques: from machine learning to meshless methods, Acta Numer. 15 (2006) 543–639.
[17] B. Shang, D.W. Apley, Full-sequential space-filling design algorithms for computer experiments, J. Qual. Technol. 53 (2) (2020) 173–196.
[18] M.C. Shewry, H.P. Wynn, Maximum entropy sampling, Appl. Stat. 14 (1987) 165–170.
[19] M.L. Stein, Interpolation of Spatial Data: Some Theory for Kriging, Springer, 1999.
[20] H. Wendland, Scattered Data Approximation, Cambridge University Press, 2005.
[21] T. Wenzel, G. Santin, B. Haasdonk, A novel class of stabilized greedy kernel approximation algorithms: Convergence, stability and uniform point distribution, J. Approx. Theory 262 (2021) 105508.

[1] https://sdb3.i3s.unice.fr/anrindex/