Design of Computer Experiments — (1) without model —

Luc Pronzato

Université Côte d'Azur, CNRS, I3S, France

Objectives

Computer experiments: based on simulations

➤ Usually, $\mathbf{x} \in \mathbb{R}^d$ ➠ observation $Y(\mathbf{x})$ (physical experiment)
➤ here, numerical simulation: $Y(\mathbf{x}) = f(\mathbf{x})$, observation = evaluation of an unknown function $f(\cdot)$
(no measurement error)

Objectives

Computer experiments: based on simulations

➤ Usually, $\mathbf{x} \in \mathbb{R}^d$ ➠ observation $Y(\mathbf{x})$ (physical experiment)
➤ here, numerical simulation: $Y(\mathbf{x}) = f(\mathbf{x})$, observation = evaluation of an unknown function $f(\cdot)$
(no measurement error)

from pairs $(\mathbf{x}_i, f(\mathbf{x}_i))$, $i = 1, 2, \ldots, n$

- optimization: find $\mathbf{x}^* = \arg\max_{\mathbf{x} \in \mathscr{X}} f(\mathbf{x})$
- inversion: construct $\{\mathbf{x} \in \mathscr{X} : f(\mathbf{x}) = T\}$
- estimation of a probability of failure: $\text{Prob}\{f(\mathbf{x}) > C\}$ when $\mathbf{x} \sim$ probability density $\varphi(\cdot)$
- sensitivity analysis
- **approximation/interpolation of $f(\cdot)$** by a predictor $\eta_n(\cdot)$, to be constructed

Objective = approximation/interpolation

$f(\mathbf{x})$ an unknown function, defined on $\mathscr{X} \subset \mathbb{R}^d$ (compact)
construct a "good" approximation $\eta_n(\cdot)$ of $f(\cdot)$ over $\mathscr{X}$ from pairs $(\mathbf{x}_i, f(\mathbf{x}_i))$, $i = 1, 2, \ldots, n$ (n not necessarily fixed *a priori*)

Objective = approximation/interpolation

$f(\mathbf{x})$ an unknown function, defined on $\mathscr{X} \subset \mathbb{R}^d$ (compact)
construct a "good" approximation $\eta_n(\cdot)$ of $f(\cdot)$ over $\mathscr{X}$ from pairs $(\mathbf{x}_i, f(\mathbf{x}_i))$, $i = 1, 2, \ldots, n$ (n not necessarily fixed *a priori*)

- Since $f(\cdot)$ is unknown, we must observe everywhere!
- maximize the spread of the n points $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ in $\mathscr{X}$

(uniformly seems reasonable and can be properly justified (Biedermann and Dette, 2001))

Objective = approximation/interpolation

$f(\mathbf{x})$ an unknown function, defined on $\mathscr{X} \subset \mathbb{R}^d$ (compact)
construct a "good" approximation $\eta_n(\cdot)$ of $f(\cdot)$ over $\mathscr{X}$ from pairs $(\mathbf{x}_i, f(\mathbf{x}_i))$, $i = 1, 2, \ldots, n$ (n not necessarily fixed *a priori*)

➠ Since $f(\cdot)$ is unknown, we must observe everywhere!
➠ maximize the spread of the n points $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ in $\mathscr{X}$
(uniformly seems reasonable and can be properly justified (Biedermann and Dette, 2001))

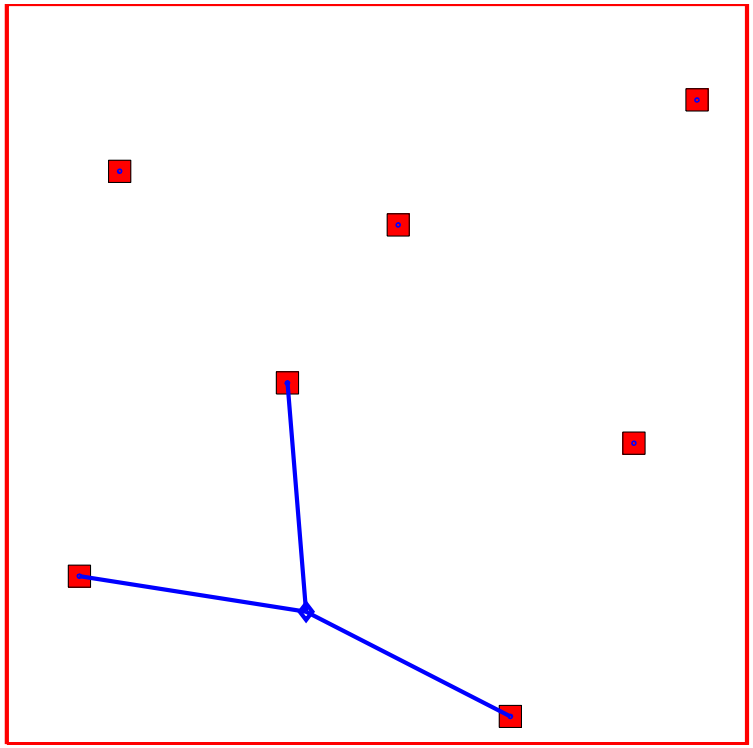
▷ $\mathbf{X}_n$ is the design (an n-point design) ◁

What does "observe everywhere" mean?

— very much based on (P., 2017)

General overview: 3 families of design criteria

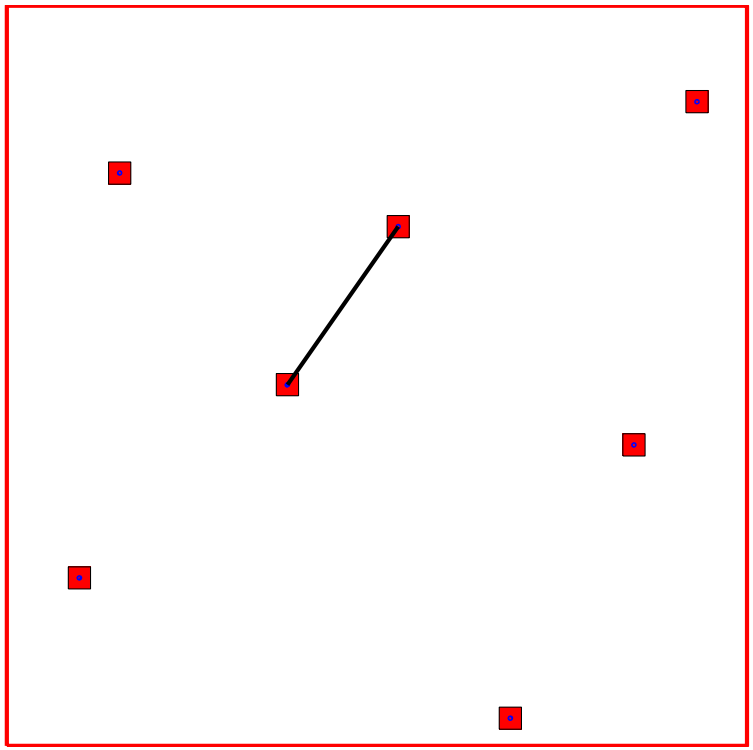
1. Inter-distance: between $\mathbf{X}_n$ and $\mathscr{X}$ (miniMax, dispersion)



. . . a bit tricky
(but we are never far from a design point)

General overview: 3 families of design criteria

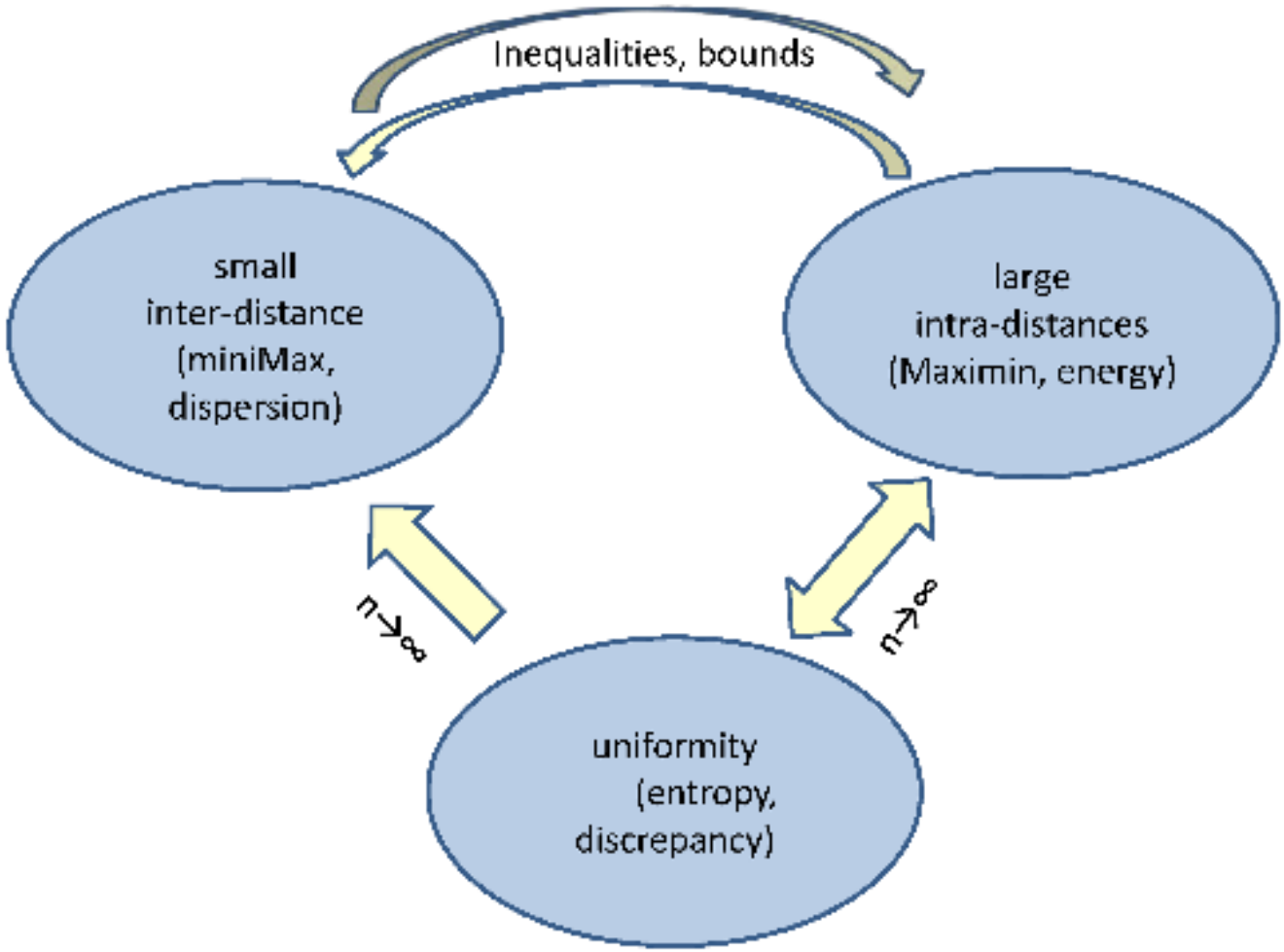
1. Inter-distance: between $\mathbf{X}_n$ and $\mathscr{X}$ (miniMax, dispersion)
2. Intra-distances: within $\mathbf{X}_n$, between design points $\mathbf{x}_i$, $i = 1, \ldots, n$
   (Maximin, energy...)



. . . easier

General overview: 3 families of design criteria

1. Inter-distance: between $\mathbf{X}_n$ and $\mathscr{X}$ (miniMax, dispersion)
2. Intra-distances: within $\mathbf{X}_n$, between design points $\mathbf{x}_i$, $i = 1, \ldots, n$
   (Maximin, energy...)
3. Uniformity of the distribution of $\mathbf{x}_i$, $i = 1, \ldots, n$
   (entropy, discrepancy ➠ generate infinite sequences of points)


Inequalities, bounds
small
inter-distance
(miniMax,
dispersion)
large
intra-distances
(Maximin, energy)
$n \to \infty$
$n \to \infty$
uniformity
(entropy,
discrepancy)

Plan

1. Geometrical space-filling criteria
   - 1.1 miniMax & Maximin: generalities
   - 1.2 Latin hypercubes
   - 1.3 miniMax (inter-distance) criterion
   - 1.4 Maximin (intra-distances) criterion
   - 1.5 Relations between Φ_{Mm} and Φ_{mM} ($d \geq 2$)
   - 1.6 Regularized Maximin, energy
2. Uniformity: quasi Monte-Carlo, discrepancy
   - 2.1 Entropy, optimal graphs
   - 2.2 Discrepancy: motivation
   - 2.3 Discrepancy criteria
   - 2.4 Low discrepancy sequences
   - 2.5 (t, m, d)-nets & (t, d)-sequences
3. Dispersion & miniMax
   - 3.1 Dispersion
   - 3.2 Low dispersion sequences
4. Conclusions part (1)

1 Geometrical space-filling criteria

1.1 miniMax & Maximin: generalities (Johnson et al., 1990)

① **miniMax**: minimize $\Phi_{mM}(\mathbf{X}_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$

$$\Phi_{mM}(\mathbf{X}_n) = d_{\text{Hausdorff}}(\mathbf{X}_n, \mathscr{X}) = \max\{\max_{\mathbf{x}\in\mathscr{X}} d(x, \mathbf{X}_n), \underbrace{\max_{\mathbf{x}_i\in\mathbf{X}_n} d(\mathbf{x}_i, \mathscr{X})}_{=0 \ (\mathbf{X}_n\in\mathscr{X}^n)}\}$$

➠ **Inter-distance** between $\mathbf{X}_n$ and $\mathscr{X}$

$$d = 1 \Leftrightarrow x_i = (2i-1)/(2n),\ i = 1, \ldots, n$$
$$\Rightarrow \Phi^*_{mM,n} = 1/(2n)$$

$d > 1 \Leftrightarrow$ **sphere-covering**

② **Maximin**: maximize $\Phi_{Mm}(\mathbf{X}_n) = \min_{i \neq j} d_{ij} = \min_{i \neq j} \|\mathbf{x}_i - \mathbf{x}_j\|$

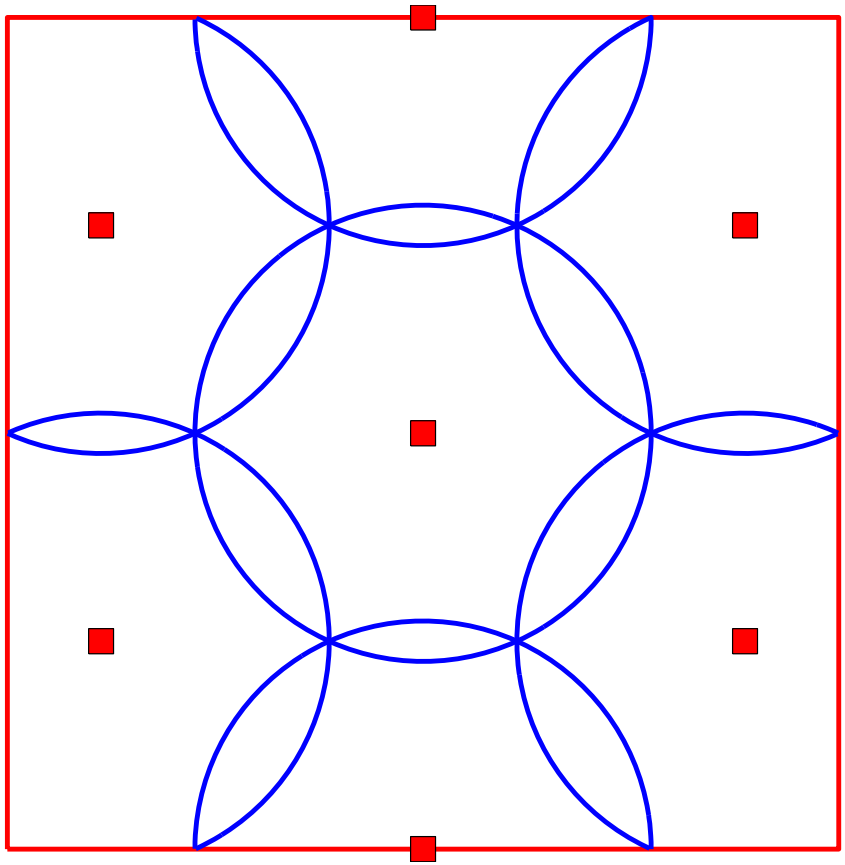
➠ **Intra-distances** for $\mathbf{X}_n$ (between points $\mathbf{x}_i$ in $\mathbf{X}_n$)

$d = 1 \Leftrightarrow x_i = (i-1)/(n-1),\ i = 1, \ldots, n$
$\Rightarrow \Phi^*_{Mm,n} = 1/(n-1)$

$d > 1 \Leftrightarrow$ **sphere-packing**

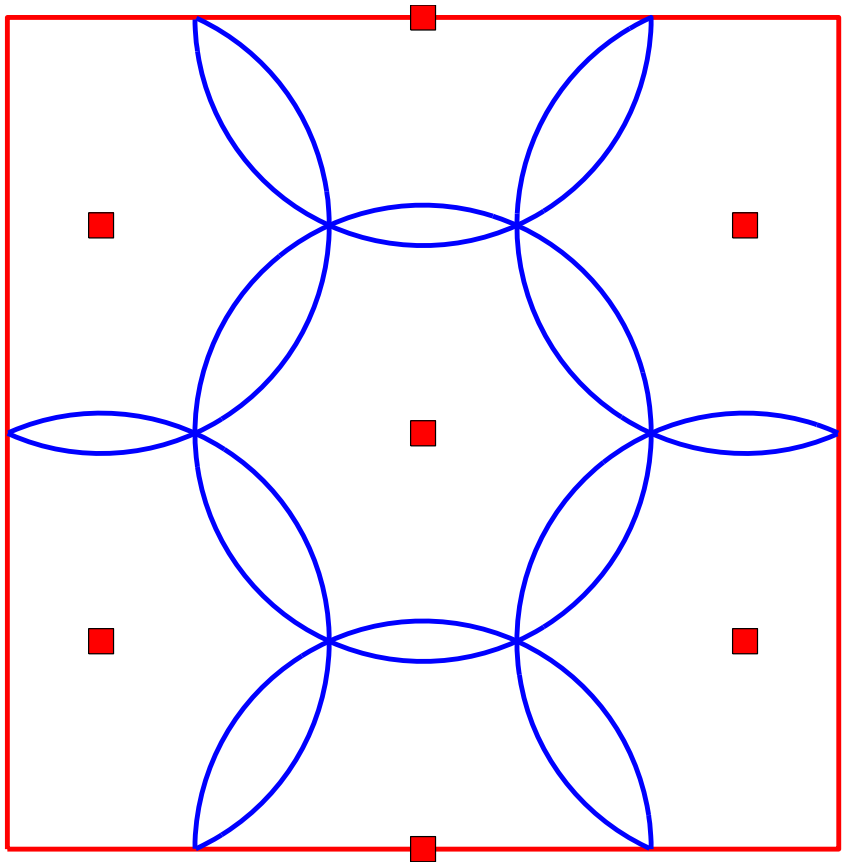
Examples :

① miniMax $d = 2, n = 7$
(radius=$\phi_{mM}(\mathbf{X}_n)$)

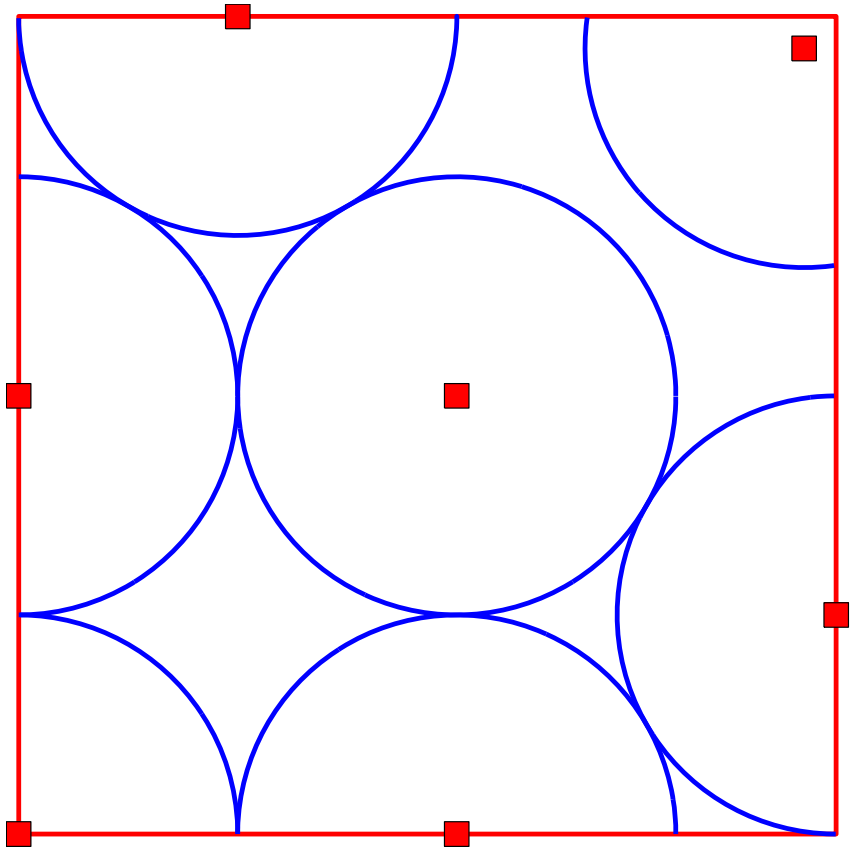


Examples :

① miniMax $d = 2, n = 7$
(radius=$\phi_{mM}(\mathbf{X}_n)$)

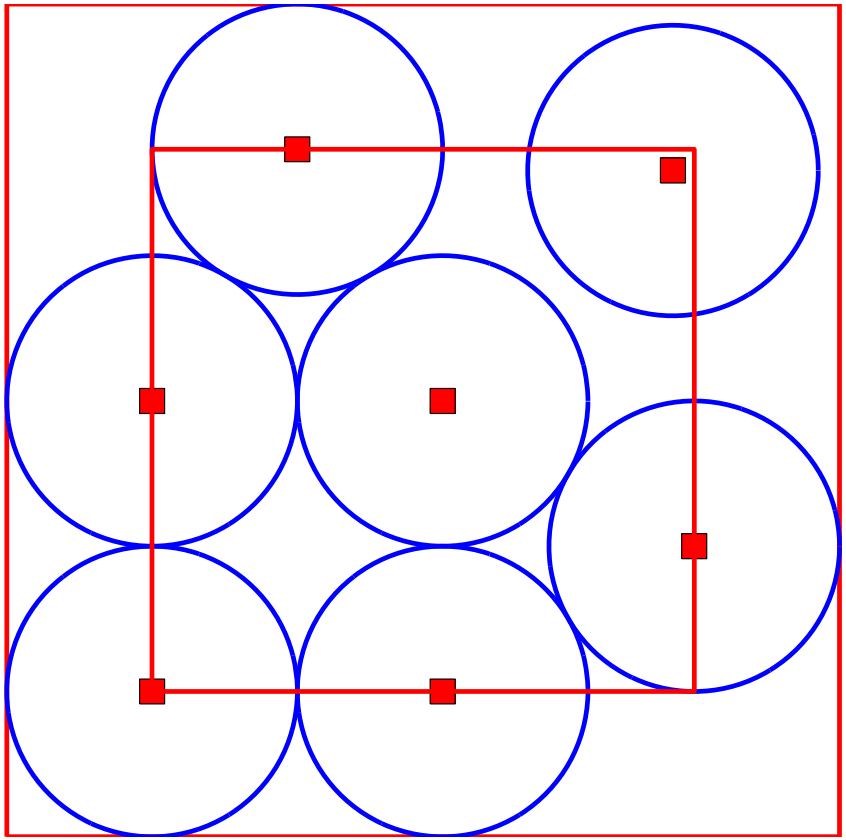


② Maximin $d = 2, n = 7$
(radius=$\phi_{Mm}(\mathbf{X}_n)/2$)



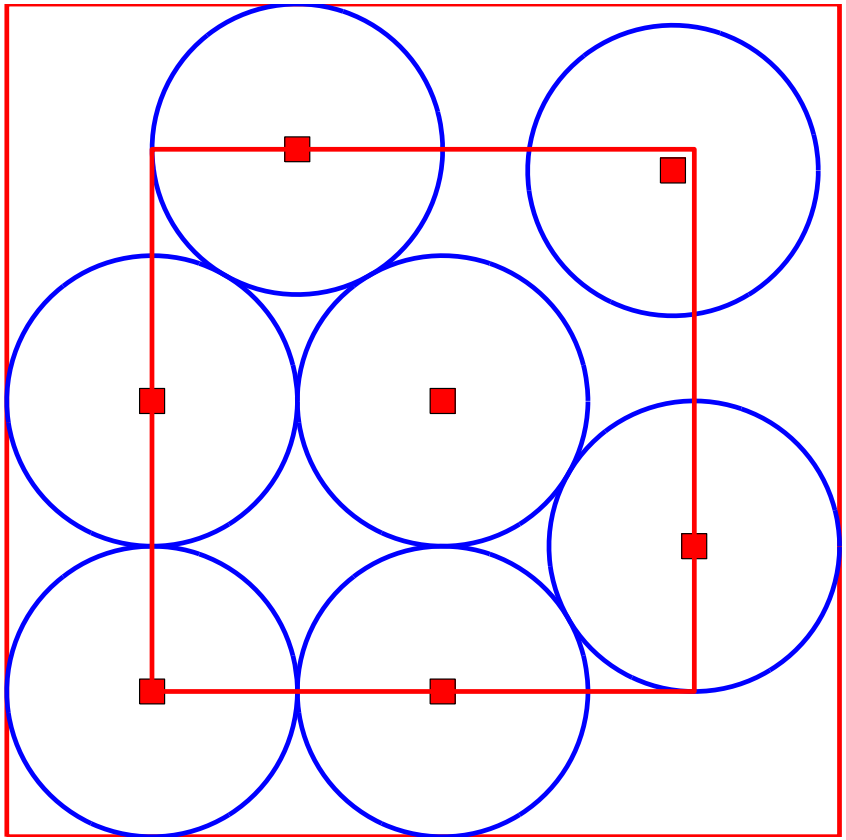
Why Maximin $\Leftrightarrow$ sphere-packing?

$d = 2, n = 7$

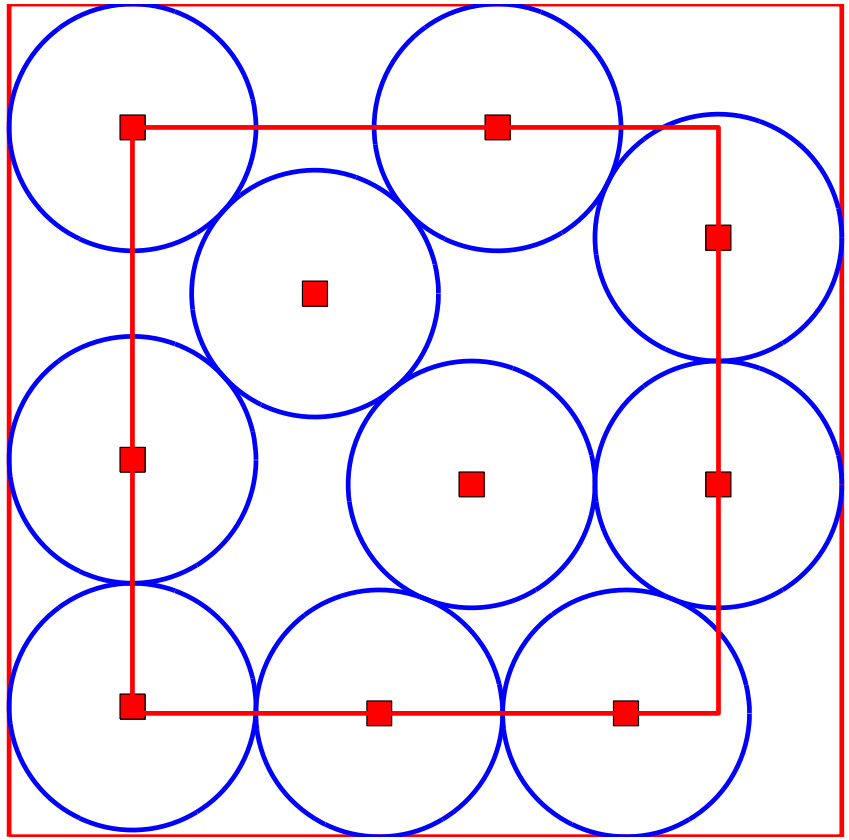


Why Maximin $\Leftrightarrow$ sphere-packing?

$d = 2, n = 7$



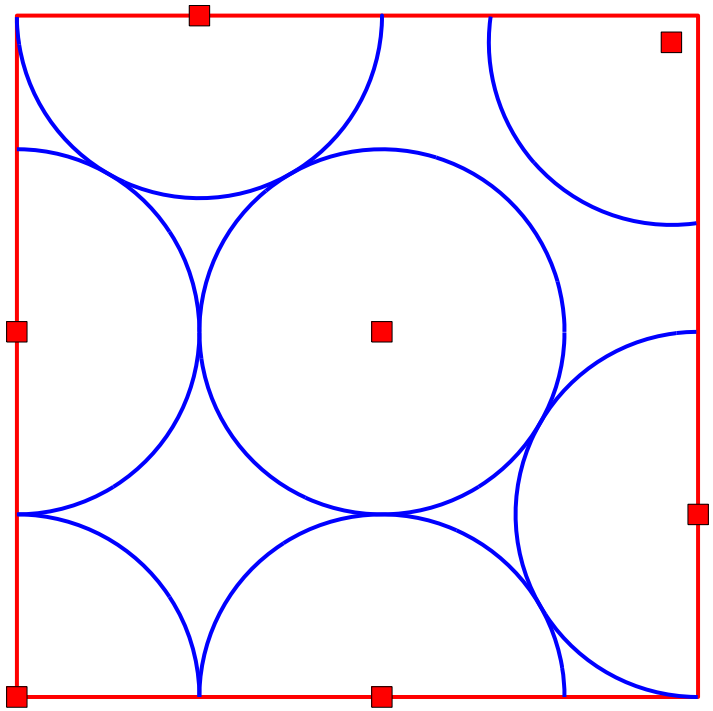
$d = 2, n = 10$



A few difficulties:

a) Local optima

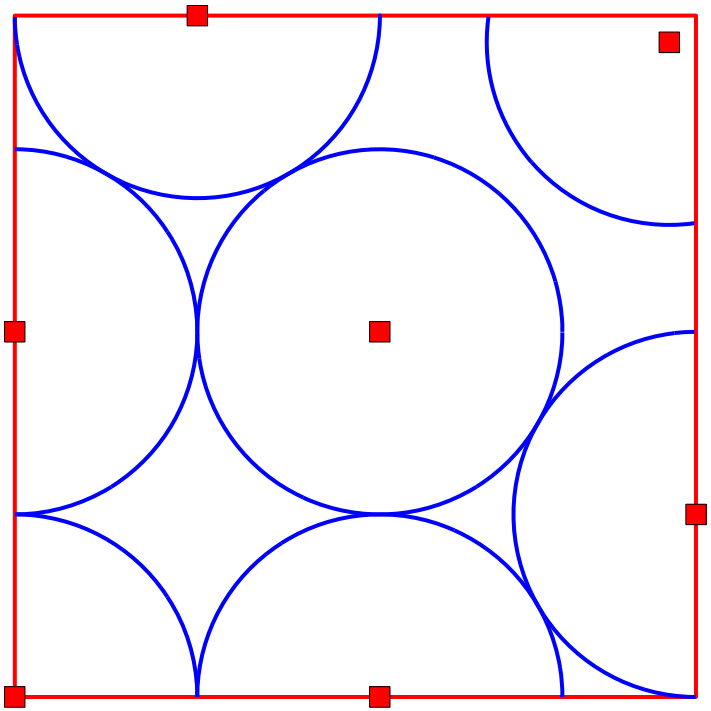
Maximin $d = 2, n = 7$, global opt.



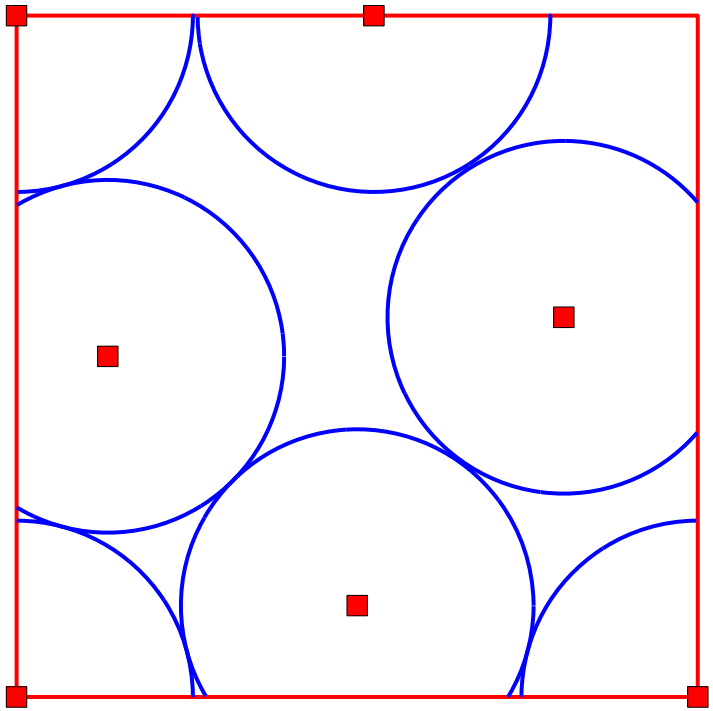
A few difficulties:

a) Local optima

Maximin $d = 2, n = 7$, global opt.



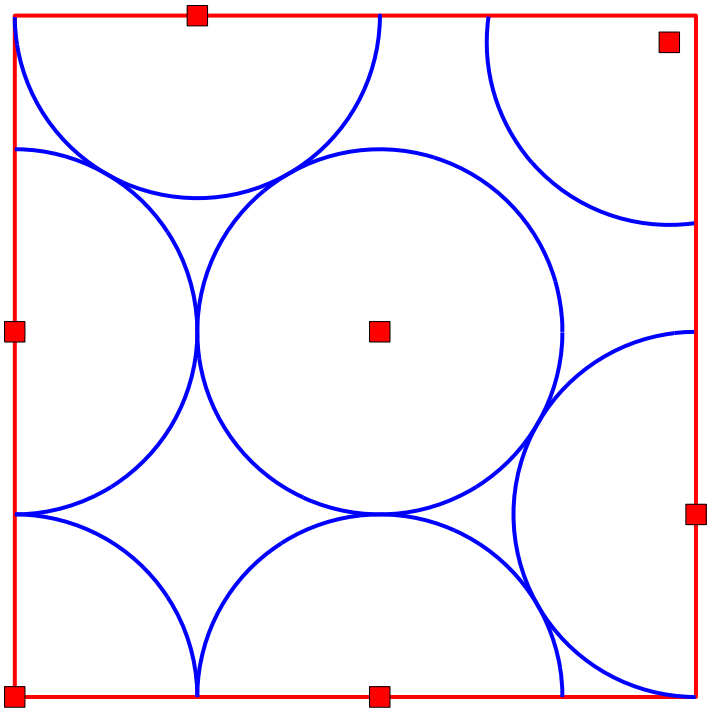
Maximin $d = 2, n = 7$, local opt.



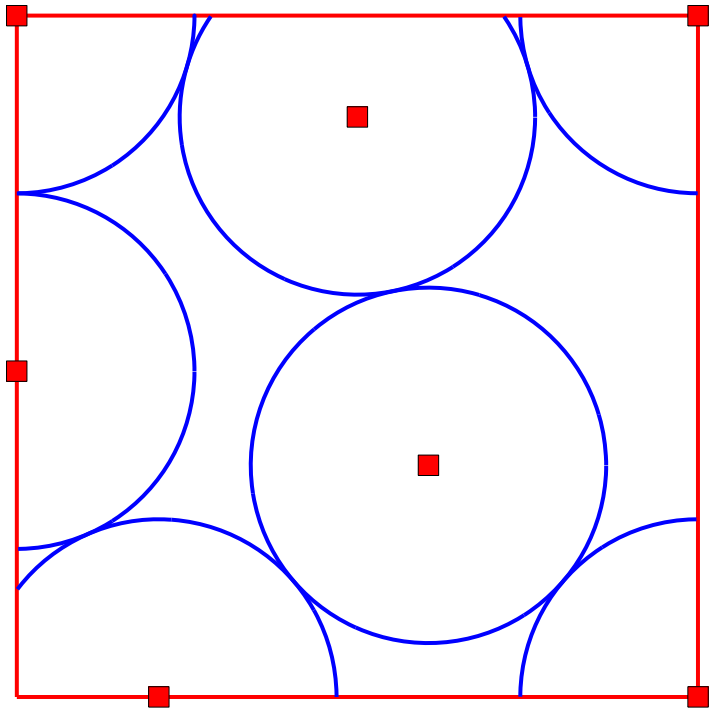
A few difficulties:

a) Local optima

Maximin $d = 2, n = 7$, global opt.



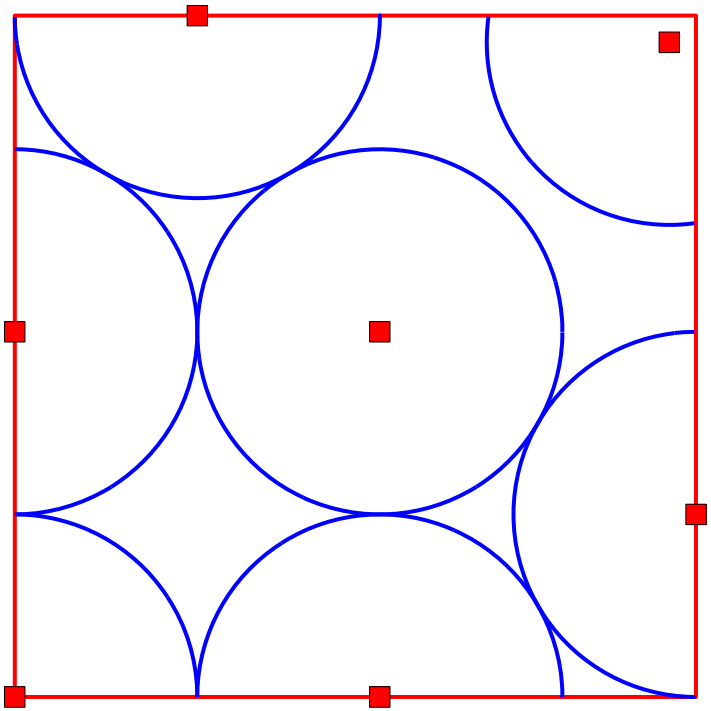
Maximin $d = 2, n = 7$, local opt.



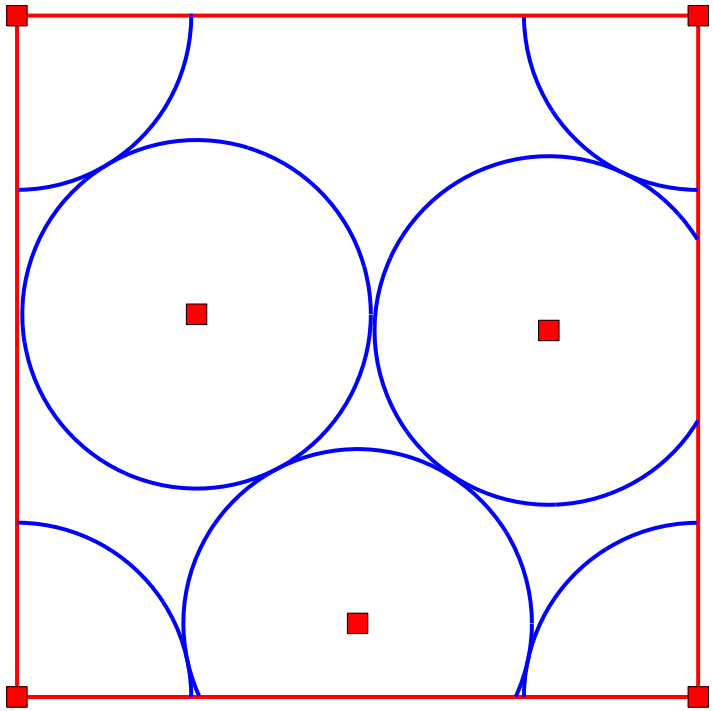
A few difficulties:

a) Local optima

Maximin $d = 2, n = 7$, global opt.



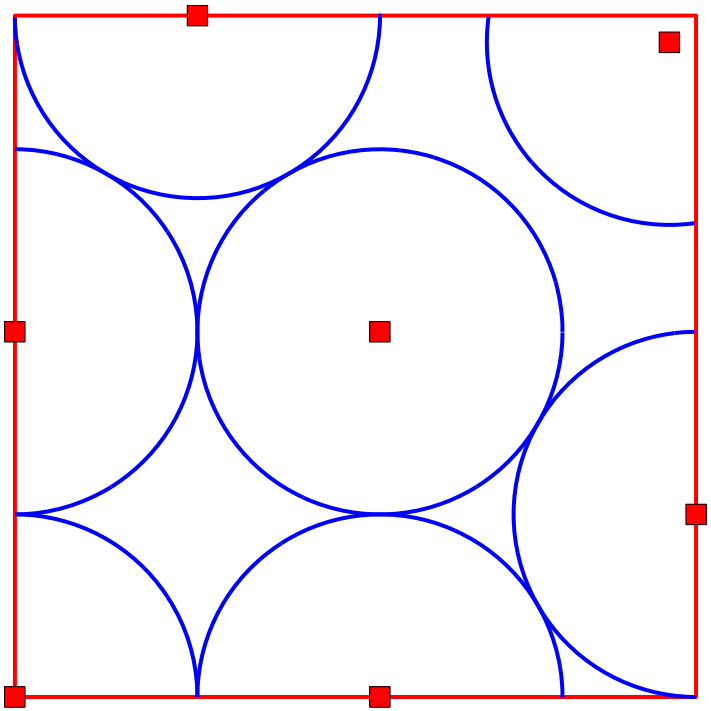
Maximin $d = 2, n = 7$, local opt.



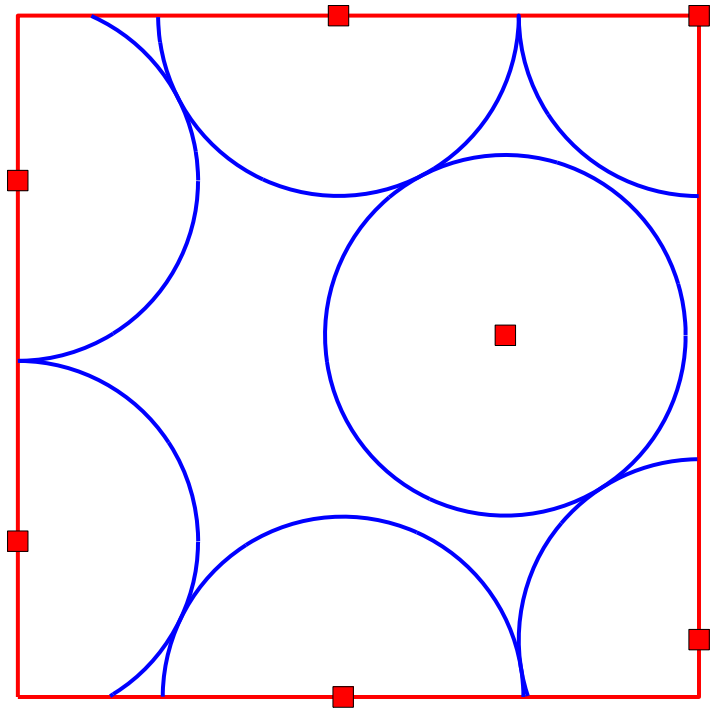
A few difficulties:

a) Local optima

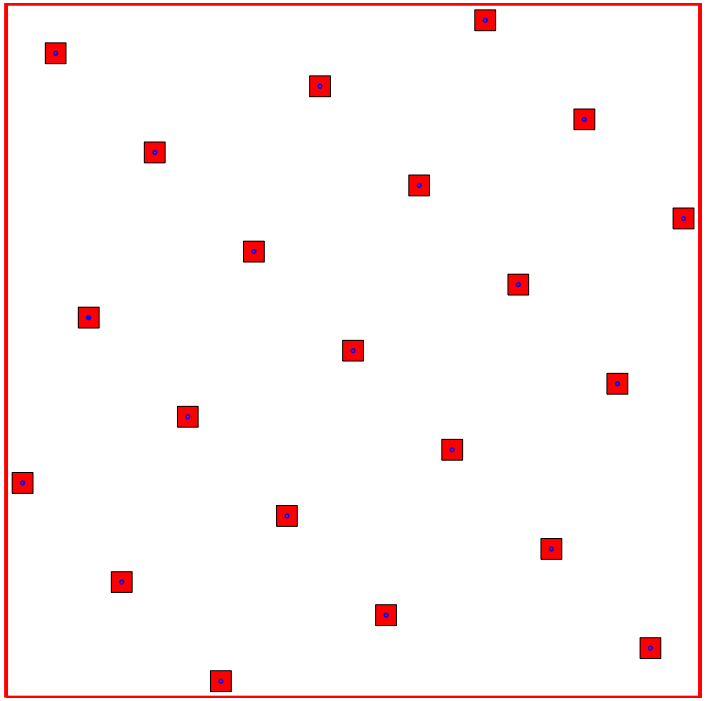
Maximin $d = 2, n = 7$, global opt.

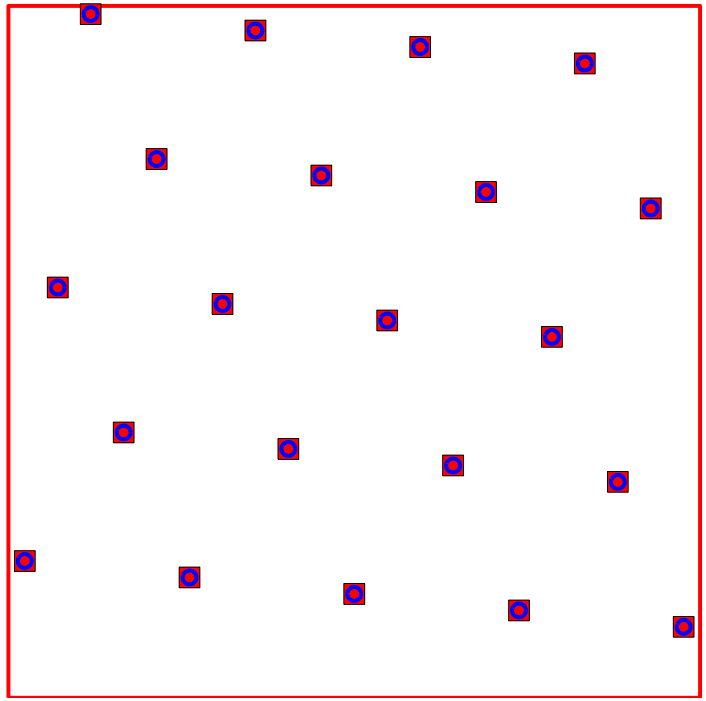


Maximin $d = 2, n = 7$, local opt.



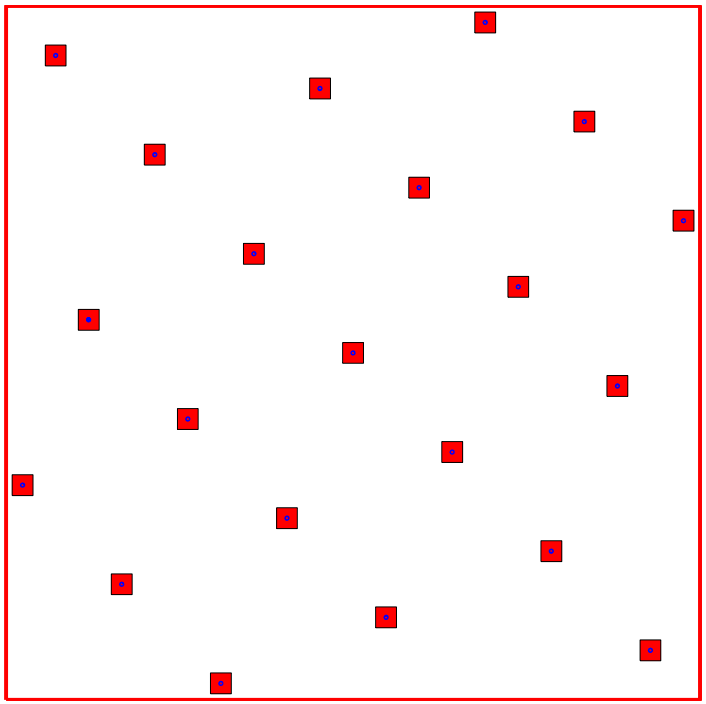
b) Misleading intuition



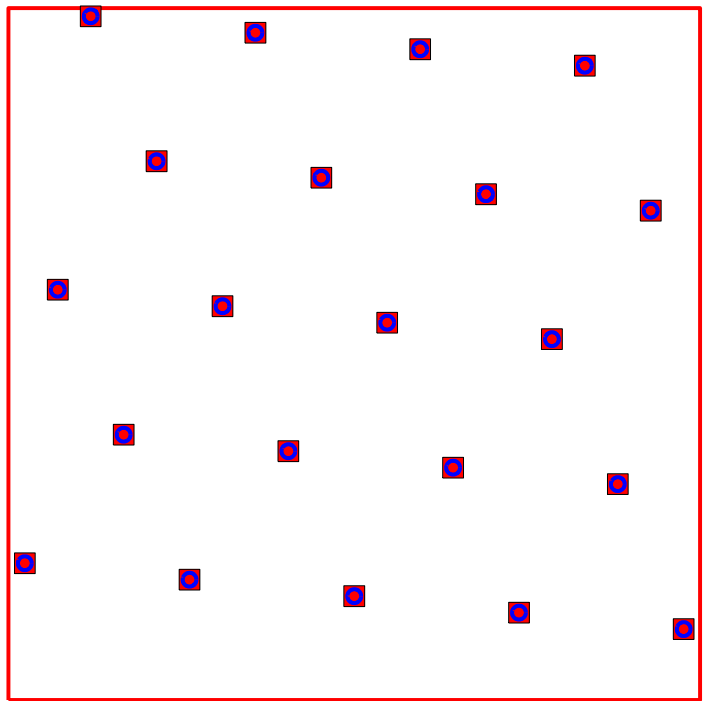


Which one is better?

b) Misleading intuition



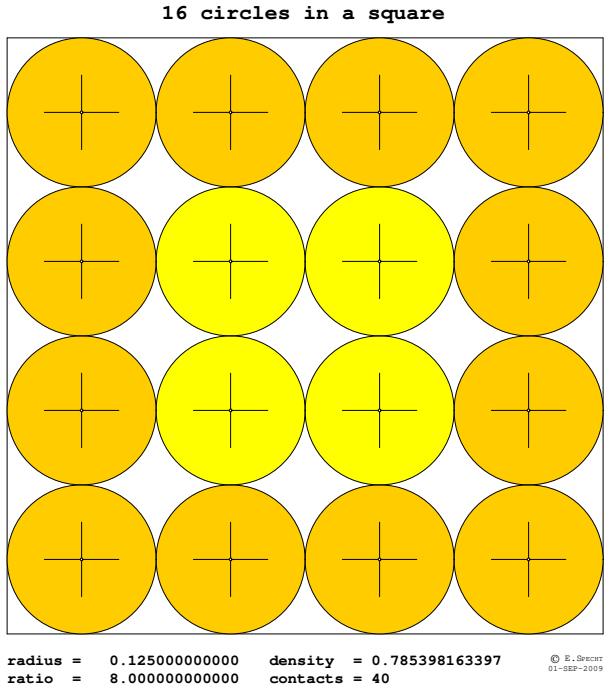
Geometry : $\Phi_{Mm}(\mathbf{X}_n) = 0.2020$
$\Phi_{mM}(\mathbf{X}_n) = 0.2357$
Uniformity : $D_{Cent,L_2}(\mathbf{X}_n) = 0.0280$
$D_{WA,L_2}(\mathbf{X}_n) = 0.0388$



Geometry : $\Phi_{Mm}(\mathbf{X}_n) = 0.2302$
$\Phi_{mM}(\mathbf{X}_n) = 0.2217$
Uniformity : $D_{Cent,L_2}(\mathbf{X}_n) = 0.0536$
$D_{WA,L_2}(\mathbf{X}_n) = 0.0633$

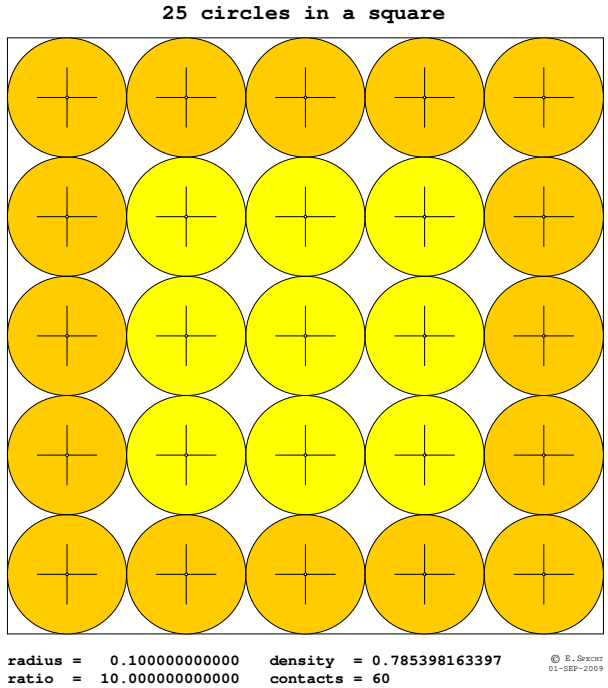
c) Sphere-packing: no trivial solution,
see `http://www.packomania.com/`

$$d = 2, n = 16$$



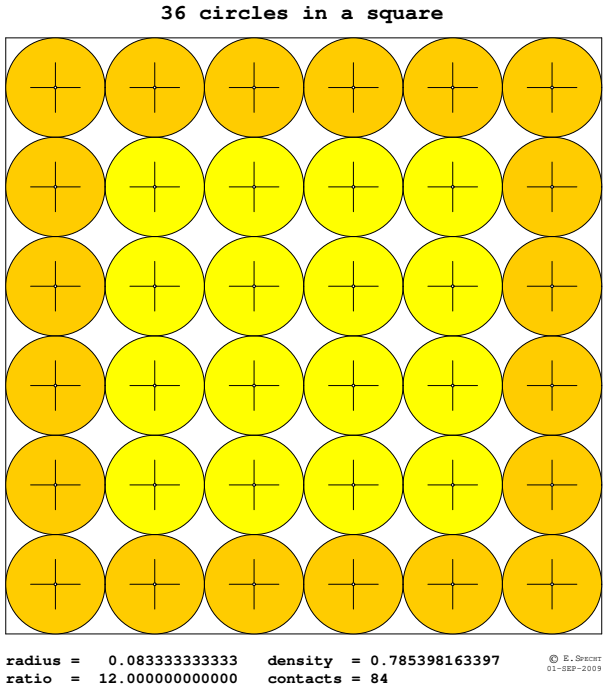
c) Sphere-packing: no trivial solution,
see `http://www.packomania.com/`

$$d = 2, n = 25$$



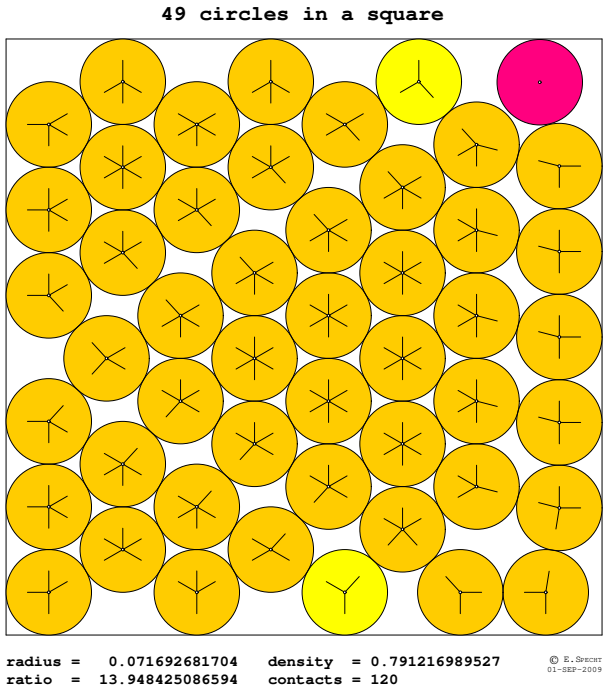
c) Sphere-packing: no trivial solution,
see `http://www.packomania.com/`

$$d = 2, n = 36$$



c) Sphere-packing: no trivial solution,
see `http://www.packomania.com/`

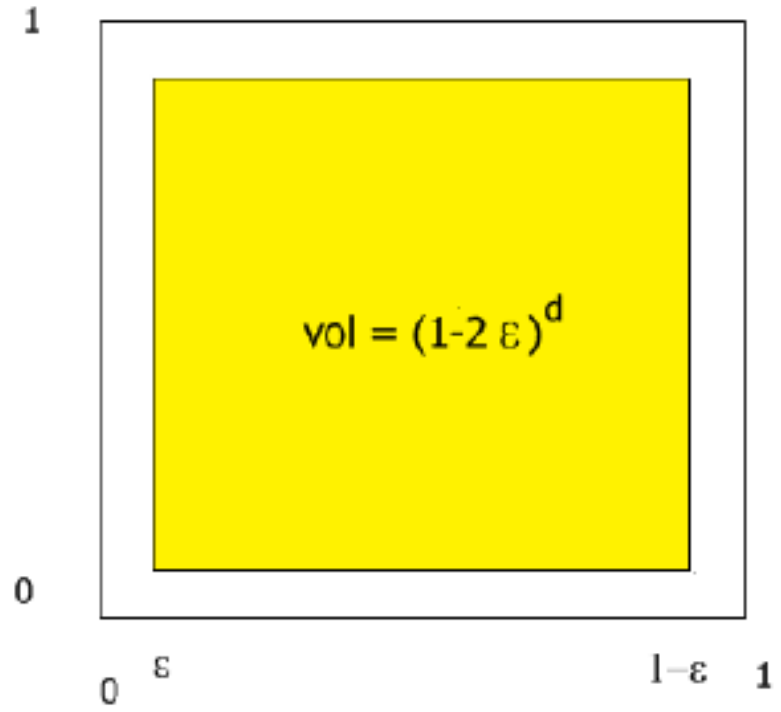
$$d = 2, n = 49$$



Cube packing is much easier! (see § 3.1)

d) Curse of dimensionality: when $d \to \infty$, all volume of $[0,1]^d$ is along the boundary

$$\mathcal{X} = \mathbb{I}_d \triangleq [0,1]^d$$



For ϵ given, volume of central part $= (1-2\epsilon)^d \to 0$ when $d \to \infty$

Ex: 64 balls in a cube, in 4 regular layers of 16 balls

$\to$ 56 touch the boundary!

e) Another issue related to dimension: big difference between a cube and a ball for large d (see (Blum et al., 2016, Chap. 2))

➤ Unit cube $\mathcal{K}_d(\mathbf{0}, 1) = [-1/2, 1/2]^d$: volume $= 1$, max distance between 2 points $= \sqrt{d}$

➤ Unit ball $\mathscr{B}_d(\mathbf{0}, 1)$: volume $= \pi^{d/2}/\Gamma(d/2+1) \to 0$ (quickly) as $d \to \infty$, max distance between 2 points $= 2$

e) Another issue related to dimension: big difference between a cube and a ball for large d (see (Blum et al., 2016, Chap. 2))

➤ Unit cube $\mathcal{K}_d(\mathbf{0}, 1) = [-1/2, 1/2]^d$: volume $= 1$, max distance between 2 points $= \sqrt{d}$

➤ Unit ball $\mathscr{B}_d(\mathbf{0}, 1)$: volume $= \pi^{d/2}/\Gamma(d/2+1) \to 0$ (quickly) as $d \to \infty$, max distance between 2 points $= 2$

$\mathcal{K}_d(\mathbf{0}, 1) \subset \mathscr{B}_d(\mathbf{0}, 1)$ for $d \leq 4$, but the vertices of the cube (at distance $\sqrt{d}/2$ from $\mathbf{0}$) lie outside $\mathscr{B}_d(\mathbf{0}, 1)$ for $d \geq 5$

e) Another issue related to dimension: big difference between a cube and a ball for large d (see (Blum et al., 2016, Chap. 2))

➤ Unit cube $\mathcal{K}_d(\mathbf{0}, 1) = [-1/2, 1/2]^d$: volume $= 1$, max distance between 2 points $= \sqrt{d}$

➤ Unit ball $\mathscr{B}_d(\mathbf{0}, 1)$: volume $= \pi^{d/2}/\Gamma(d/2+1) \to 0$ (quickly) as $d \to \infty$, max distance between 2 points $= 2$

$\mathcal{K}_d(\mathbf{0}, 1) \subset \mathscr{B}_d(\mathbf{0}, 1)$ for $d \leq 4$, but the vertices of the cube (at distance $\sqrt{d}/2$ from $\mathbf{0}$) lie outside $\mathscr{B}_d(\mathbf{0}, 1)$ for $d \geq 5$

$\text{vol}[\mathcal{K}_d(\mathbf{0}, 1) \setminus \mathscr{B}_(\mathbf{0}, 1)] \to 1$ as $d \to \infty$, but $\mathscr{B}_d(\mathbf{0}, 1) \not\subset \mathcal{K}_d(\mathbf{0}, 1)$! (the centers of faces are always at distance $1/2$ from$\mathbf{0}$)

$\Longrightarrow$ For large d, working within a cube is much more difficult than working within a ball

e) Another issue related to dimension: big difference between a cube and a ball for large d (see (Blum et al., 2016, Chap. 2))

➤ Unit cube $\mathcal{K}_d(\mathbf{0}, 1) = [-1/2, 1/2]^d$: volume $= 1$, max distance between 2 points $= \sqrt{d}$

➤ Unit ball $\mathscr{B}_d(\mathbf{0}, 1)$: volume $= \pi^{d/2}/\Gamma(d/2+1) \to 0$ (quickly) as $d \to \infty$, max distance between 2 points $= 2$

$\mathcal{K}_d(\mathbf{0}, 1) \subset \mathscr{B}_d(\mathbf{0}, 1)$ for $d \leq 4$, but the vertices of the cube (at distance $\sqrt{d}/2$ from $\mathbf{0}$) lie outside $\mathscr{B}_d(\mathbf{0}, 1)$ for $d \geq 5$

$\text{vol}[\mathcal{K}_d(\mathbf{0}, 1) \setminus \mathscr{B}_(\mathbf{0}, 1)] \to 1$ as $d \to \infty$, but $\mathscr{B}_d(\mathbf{0}, 1) \not\subset \mathcal{K}_d(\mathbf{0}, 1)$! (the centers of faces are always at distance $1/2$ from$\mathbf{0}$)

$\Longrightarrow$ For large d, working within a cube is much more difficult than working within a ball

We shall lower our ambitions: finding an optimal design is extremely difficult, we shall only try to find "reasonable" designs

1.2 Latin hypercubes

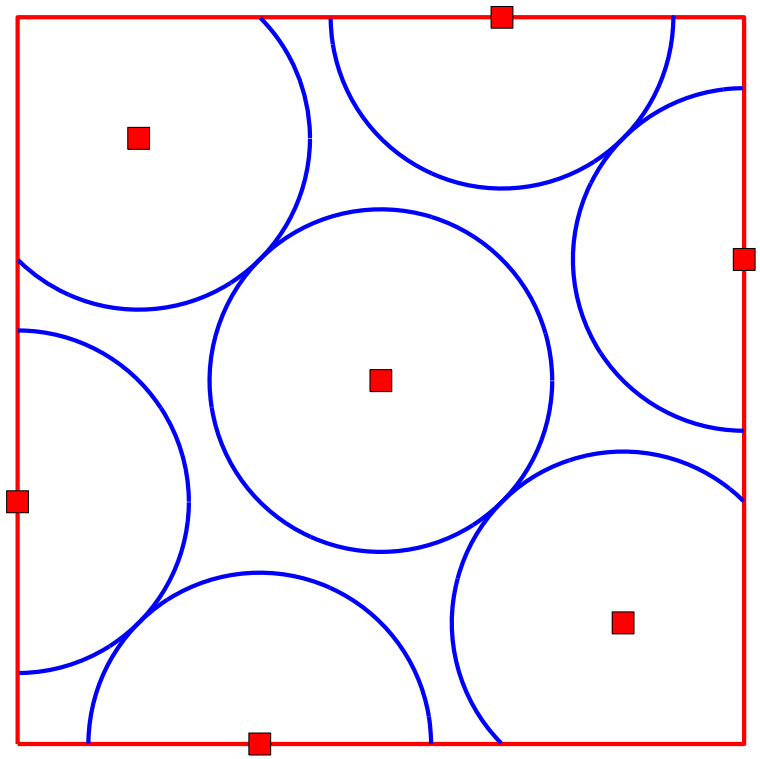
Objective: ensure good projection properties along each principal axis

each 1d projection is Maximin-optimal

$\{\mathbf{x}_i\}_\ell \in \{0, \frac{1}{n-1}, \ldots, \frac{k-1}{n-1}, \ldots, 1\}$ for all $\ell = 1, \ldots, d$

➠ only $(n!)^{d-1}$ possible designs

Maximin-optimal Lh ($d = 2, n = 7$, radius=$\phi_{Mm}(\mathbf{X}_n)/2$)



1.2 Latin hypercubes

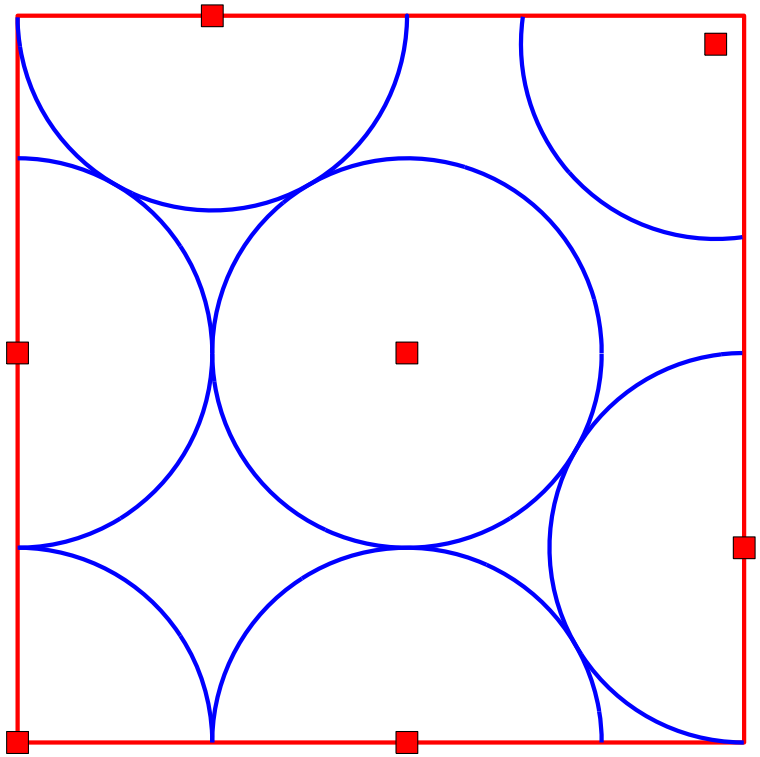
Objective: ensure good projection properties along each principal axis

each 1d projection is Maximin-optimal

$\{\mathbf{x}_i\}_\ell \in \{0, \frac{1}{n-1}, \ldots, \frac{k-1}{n-1}, \ldots, 1\}$ for all $\ell = 1, \ldots, d$

➠ only $(n!)^{d-1}$ possible designs

Maximin-optimal, not Lh, ($d = 2, n = 7$, radius=$\phi_{Mm}(\mathbf{X}_n)/2$)



1.2 Latin hypercubes

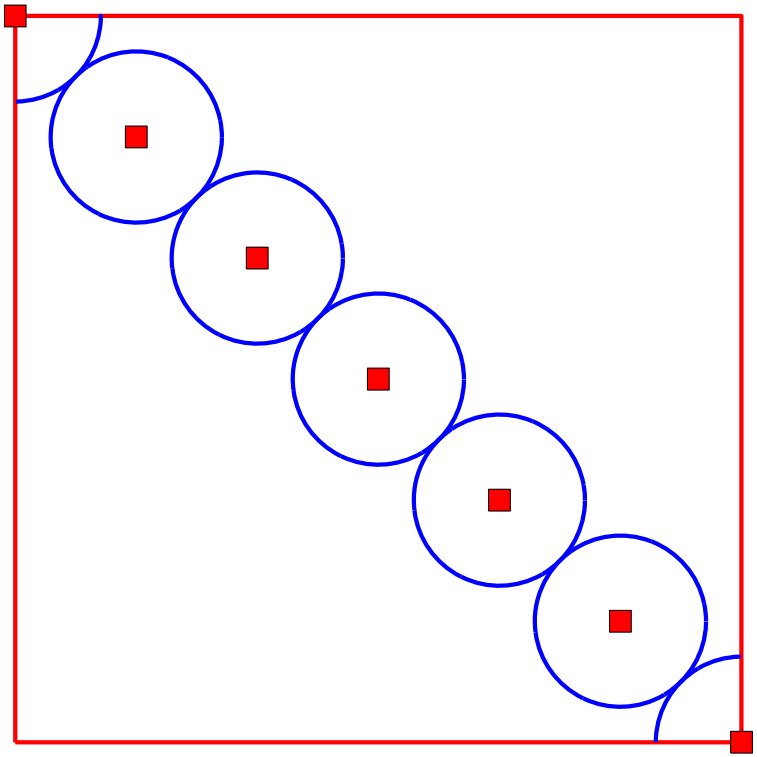
Objective: ensure good projection properties along each principal axis

each 1d projection is Maximin-optimal

$\{\mathbf{x}_i\}_\ell \in \{0, \frac{1}{n-1}, \ldots, \frac{k-1}{n-1}, \ldots, 1\}$ for all $\ell = 1, \ldots, d$

➠ only $(n!)^{d-1}$ possible designs

Lh, not Maximin-optimal, ($d = 2, n = 7$, radius=$\phi_{Mm}(\mathbf{X}_n)/2$)



The Lh property only ensures good 1d projection properties!

➠ we need to optimize another space-filling criterion (in dimension d)
(typically, using simulated annealing, other heuristics may be considered)

The Lh property only ensures good 1d projection properties!

➠ we need to optimize another space-filling criterion (in dimension d)
(typically, using simulated annealing, other heuristics may be considered)

➠ Optimizing within the class of Lh designs ensures good 1d projection properties

Important when $f(\cdot)$ may possibly not depend on some input factors $\{\mathbf{x}\}_\ell$:

➤ no repetition of points a factor is removed
➤ the projection on $d' < d$ components is still a Lh
(but not necessarily with a good distribution of points if $d' > 1$)

Abundant literature since (McKay et al., 1979), see (Viana, 2013)
The Lh constraint worsens the space-filling property

⏩ miniMax

Optimization within the class of Lh designs:

Let $\mathbf{X}_n$ be a n-point Lh design:

- choose a coordinate ℓ (among d)
- choose a pair of points $\mathbf{x}_i$ and $\mathbf{x}_j$ in $\mathbf{X}_n$
- exchange their ℓ-th coordinate
- ➠ $\mathbf{X}_n^+$, which is still a Lh design ($dn(n-1)/2$ possible constructions)

(one may also exchange several pairs of points simultaneously)

Optimization within the class of Lh designs:

Let $\mathbf{X}_n$ be a n-point Lh design:

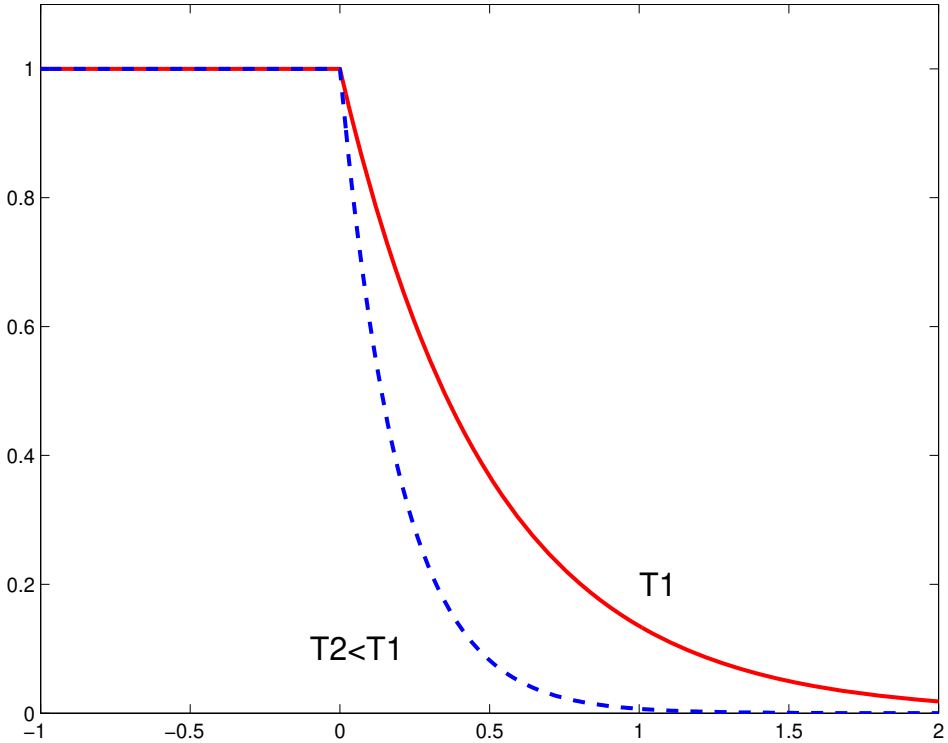
- choose a coordinate ℓ (among d)
- choose a pair of points $\mathbf{x}_i$ and $\mathbf{x}_j$ in $\mathbf{X}_n$
- exchange their ℓ-th coordinate
- ➠ $\mathbf{X}_n^+$, which is still a Lh design ($dn(n-1)/2$ possible constructions)

(one may also exchange several pairs of points simultaneously)

Simulated annealing (minimization of $\Phi(\cdot)$) — principle:

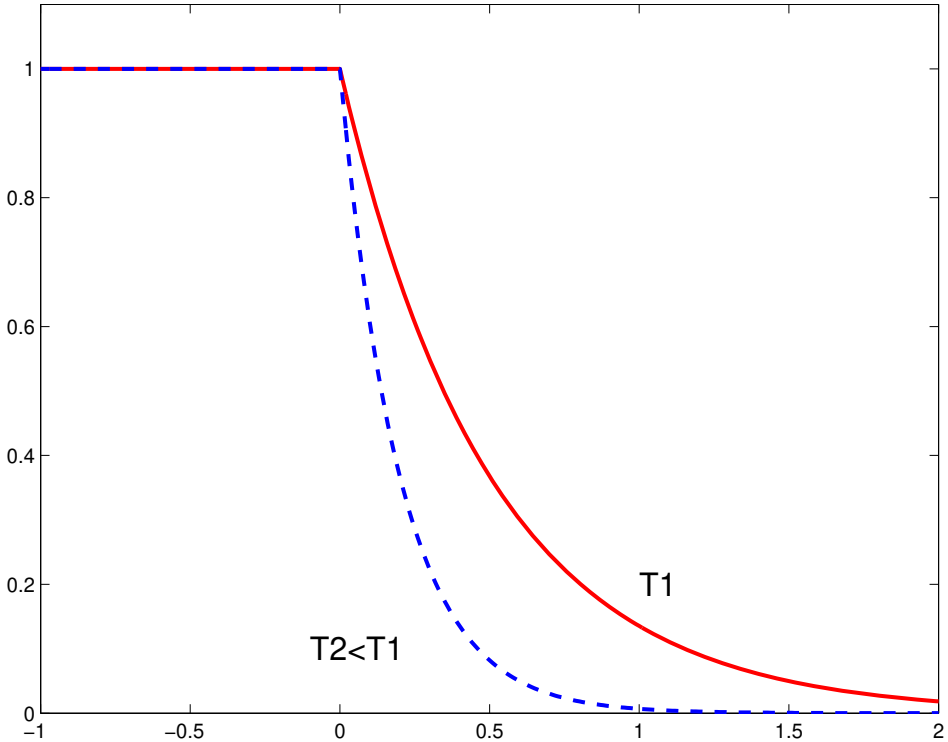
- 0) start from a Lh design $\mathbf{X}_n^0$, set $k = 0$
- 1) generate a Lh design $\mathbf{X}_n^{k+}$ from $\mathbf{X}_n^k$
- 2) calculate $\Delta\Phi_k = \Phi(\mathbf{X}_n^{k+}) - \Phi(\mathbf{X}_n^k)$
- 3) Accept $\mathbf{X}_n^{k+}$, i.e., do $\mathbf{X}_n^{k+1} = \mathbf{X}_n^{k+}$ with probability $P_k = \min\left\{1, \exp\left(-\frac{\Delta\Phi_k}{T_k}\right)\right\}$, keep $\mathbf{X}_n^{k+1} = \mathbf{X}_n^k$ with prob. $1 - P_k$
  $k \leftarrow k+1$, return to 1

$\mathbf{X}_n^{k+}$ such that $\Delta\Phi_k < 0$ is always accepted
$\mathbf{X}_n^{k+}$ such that $\Delta\Phi_k > 0$ is more often accepted for T_1 than for $T_2 < T_1$



➠ Take T_0 large enough (to escape from local optimas),
then decrease T_k (slowly enough)
(for instance, $T_k = \frac{T_0}{\log(k+1)}$,
or $T_k = \alpha^k T_0$ with $\alpha < 1$)

$\mathbf{X}_n^{k+}$ such that $\Delta\Phi_k < 0$ is always accepted
$\mathbf{X}_n^{k+}$ such that $\Delta\Phi_k > 0$ is more often accepted for T_1 than for $T_2 < T_1$



➠ Take T_0 large enough (to escape from local optimas),
then decrease T_k (slowly enough)
(for instance, $T_k = \frac{T_0}{\log(k+1)}$,
or $T_k = \alpha^k T_0$ with $\alpha < 1$)

Many variants, abundant literature...

Always store the best $\mathbf{X}_n^k$ found along the trajectory of the algorithm!!!

➠ Ensures convergence to the optimum when $k \to \infty$ under rather general conditions

1.3 miniMax criterion $\Phi_{mM}(\mathbf{X}_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$

Φ_{mM} is interesting for approximation:

Any $\mathbf{x}$ in $\mathscr{X}$ is at most at distance Φ_{mM} from a design point $\mathbf{x}_i$

Evaluation of $\Phi_{mM}(X_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_{i=1,\ldots,n} \|\mathbf{x} - \mathbf{x}_i\| = \max_{\mathbf{x}\in\mathscr{X}} d(\mathbf{x}, X_n)$:

we need to find a $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathscr{X}} d(\mathbf{x}, X_n)$

Key idea: replace $\arg\max_{\mathbf{x}\in\mathscr{X}} d(\mathbf{x}, X_n)$ by $\arg\max_{\mathbf{x}\in\mathscr{X}_Q} d(\mathbf{x}, X_n)$ for a suitable finite $\mathscr{X}_Q \subset \mathscr{X}$

1.3 miniMax criterion $\Phi_{mM}(\mathbf{X}_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$

Φ_{mM} is interesting for approximation:

Any $\mathbf{x}$ in $\mathscr{X}$ is at most at distance Φ_{mM} from a design point $\mathbf{x}_i$

Evaluation of $\Phi_{mM}(X_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_{i=1,\ldots,n} \|\mathbf{x} - \mathbf{x}_i\| = \max_{\mathbf{x}\in\mathscr{X}} d(\mathbf{x}, X_n)$:

we need to find a $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathscr{X}} d(\mathbf{x}, X_n)$

Key idea: replace $\arg\max_{\mathbf{x}\in\mathscr{X}} d(\mathbf{x}, X_n)$ by $\arg\max_{\mathbf{x}\in\mathscr{X}_Q} d(\mathbf{x}, X_n)$ for a suitable finite $\mathscr{X}_Q \subset \mathscr{X}$

0/ Usual trick: $\mathscr{X}_Q =$ **regular grid** or first Q points of a **Low Discrepancy Sequence** in $\mathscr{X}$

1.3 miniMax criterion $\Phi_{mM}(\mathbf{X}_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$

Φ_{mM} is interesting for approximation:

Any $\mathbf{x}$ in $\mathscr{X}$ is at most at distance Φ_{mM} from a design point $\mathbf{x}_i$

Evaluation of $\Phi_{mM}(X_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_{i=1,\ldots,n} \|\mathbf{x} - \mathbf{x}_i\| = \max_{\mathbf{x}\in\mathscr{X}} d(\mathbf{x}, X_n)$:

we need to find a $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathscr{X}} d(\mathbf{x}, X_n)$

Key idea: replace $\arg\max_{\mathbf{x}\in\mathscr{X}} d(\mathbf{x}, X_n)$ by $\arg\max_{\mathbf{x}\in\mathscr{X}_Q} d(\mathbf{x}, X_n)$ for a suitable finite $\mathscr{X}_Q \subset \mathscr{X}$

0/ Usual trick: $\mathscr{X}_Q =$ **regular grid** or first Q points of a **Low Discrepancy Sequence** in $\mathscr{X}$

➠ $\Phi_{mM}(X_n; \mathscr{X}_Q) \leq \Phi_{mM}(X_n)$ (optimistic result)

requires $Q = \mathcal{O}(1/\epsilon^d)$ to have $\Phi_{mM}(X_n) < \Phi_{mM}(X_n; \mathscr{X}_Q) + \epsilon$

1.3 miniMax criterion $\Phi_{mM}(\mathbf{X}_n) = \max_{\mathbf{x}\in\mathcal{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$

Φ_{mM} is interesting for approximation:

Any $\mathbf{x}$ in $\mathcal{X}$ is at most at distance Φ_{mM} from a design point $\mathbf{x}_i$

Evaluation of $\Phi_{mM}(X_n) = \max_{\mathbf{x}\in\mathcal{X}} \min_{i=1,\ldots,n} \|\mathbf{x} - \mathbf{x}_i\| = \max_{\mathbf{x}\in\mathcal{X}} d(\mathbf{x}, X_n)$:

we need to find a $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{X}} d(\mathbf{x}, X_n)$

Key idea: replace $\arg\max_{\mathbf{x}\in\mathcal{X}} d(\mathbf{x}, X_n)$ by $\arg\max_{\mathbf{x}\in\mathcal{X}_Q} d(\mathbf{x}, X_n)$ for a suitable finite $\mathcal{X}_Q \subset \mathcal{X}$

0/ Usual trick: $\mathcal{X}_Q =$ **regular grid** or first Q points of a **Low Discrepancy Sequence** in $\mathcal{X}$

➠ $\Phi_{mM}(X_n; \mathcal{X}_Q) \leq \Phi_{mM}(X_n)$ (optimistic result)

requires $Q = \mathcal{O}(1/\epsilon^d)$ to have $\Phi_{mM}(X_n) < \Phi_{mM}(X_n; \mathcal{X}_Q) + \epsilon$

A/ & B/ Tools from **algorithmic geometry** ($d \lesssim 5$) ➔ exact result through the construction of a suitable $\mathcal{X}_Q$

C/ MCMC $\mathcal{X}_Q =$ adaptive grid

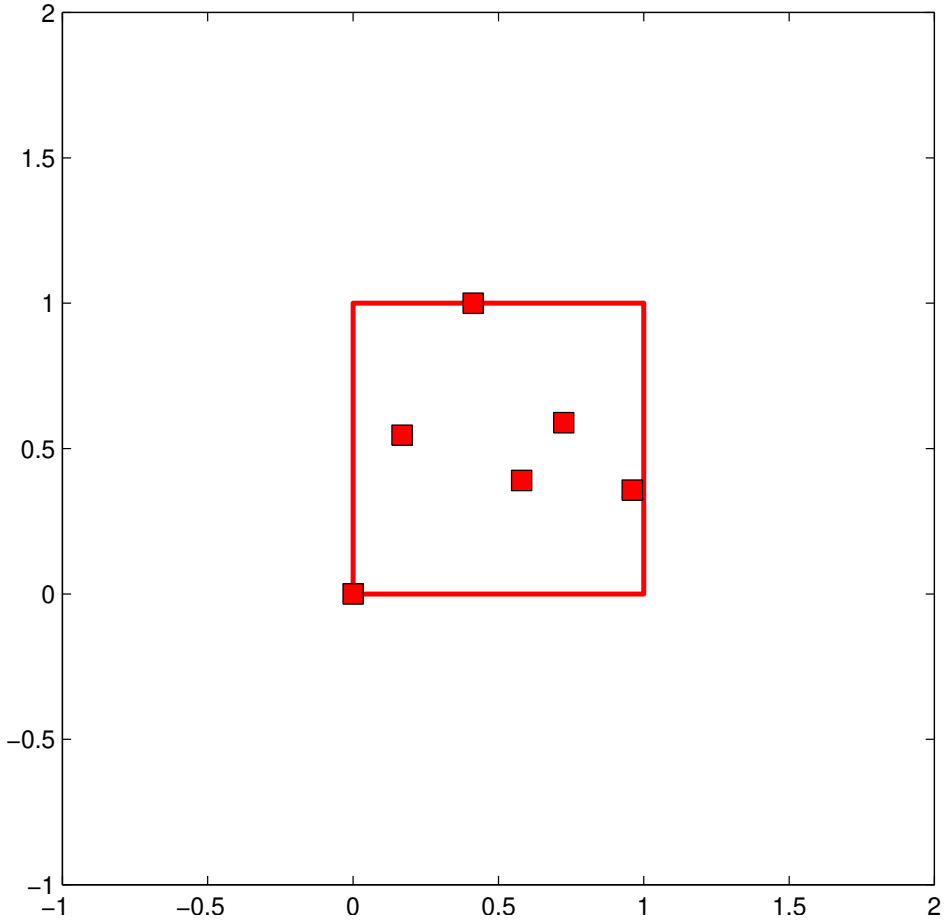
A/ Delaunay triangulation

$\mathscr{X}$ = **hypercube**, see (P. and Müller, 2012)

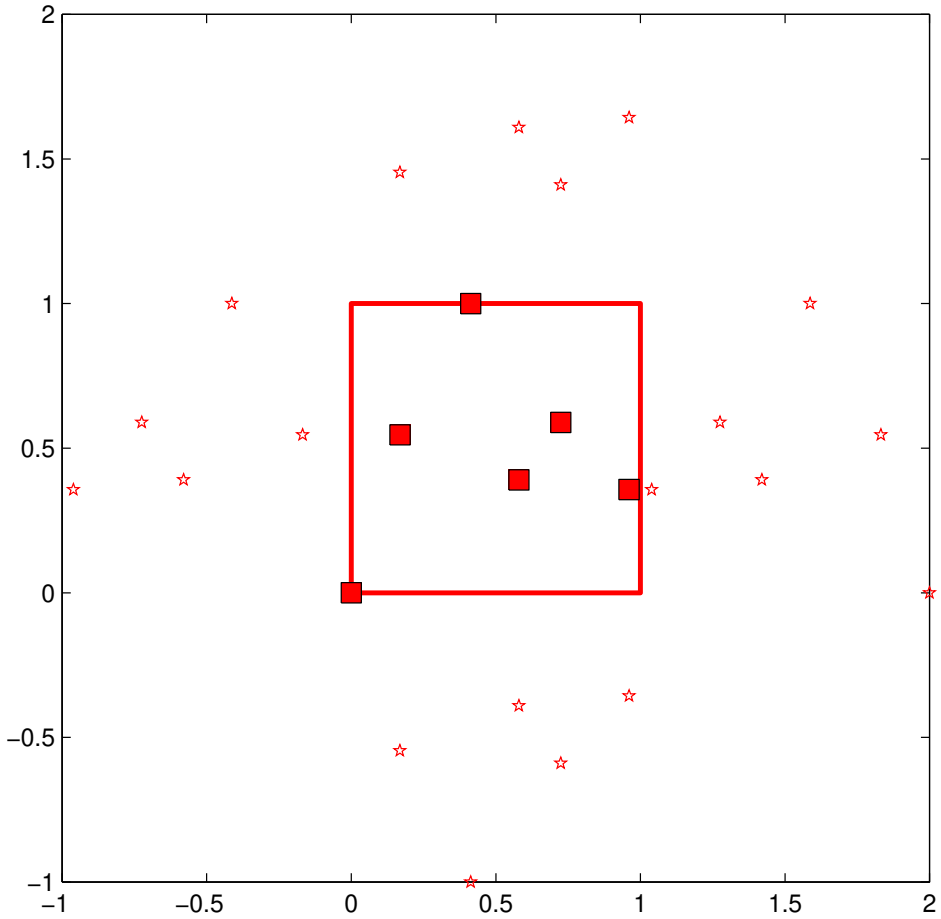
Delaunay

- $\mathbf{X}_n$ ($= n$ points in $\mathscr{X} = [0,1]^d$), consider X'_m, with $m = (2d+1)n$ points, formed by X_n and its $2d$ reflections through the $(d-1)$-dimensional faces of $\mathscr{X}$
- Compute the Delaunay triangulation of $X'_m \rightarrow$ d-dimensional simplices (each one having $d+1$ vertices), with circumscribed spheres $\mathcal{S}_j$ not containing any point of X'_m in their interior
- $\max_{\mathbf{x} \in \mathscr{X}} d(\mathbf{x}, X_n)$ is attained for $\mathbf{x}$ = centre of one $\mathcal{S}_j$
- Take $\mathscr{X}_Q$ = finite set given by centres of $\mathcal{S}_j$ that belong to $\mathscr{X}$
- $Q = |\mathscr{X}_Q| = \mathcal{O}(m^{\lceil d/2 \rceil})$, computational time $= \mathcal{O}(m^{1+\lceil d/2 \rceil}) \rightarrow$ small d only

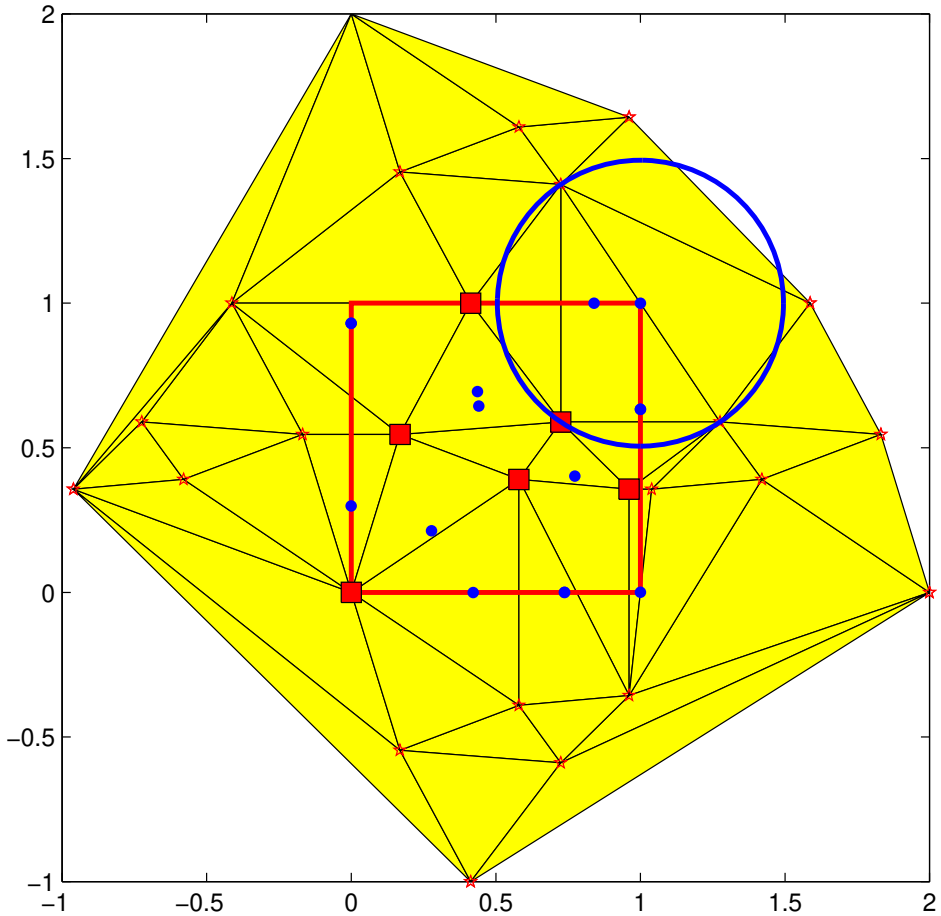
$\mathscr{X}_Q = \{$ centres of circumscribed spheres to Delaunay simplices $\}$



$\mathscr{X}_Q = \{$ centres of circumscribed spheres to Delaunay simplices $\}$



$\mathscr{X}_Q = \{$ centres of circumscribed spheres to Delaunay simplices $\}$



$n = 6$ points, 45 triangles, 12 circles (the largest one is plotted)

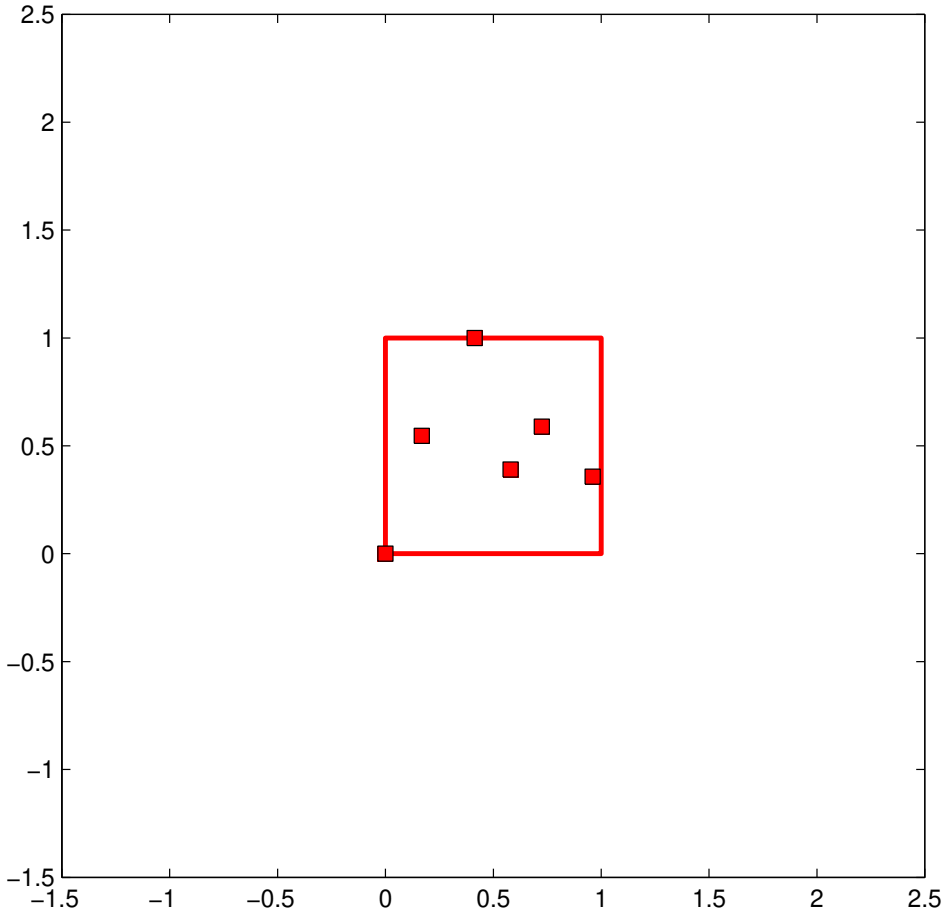
B/ Voronoï tessellation

Voronoï

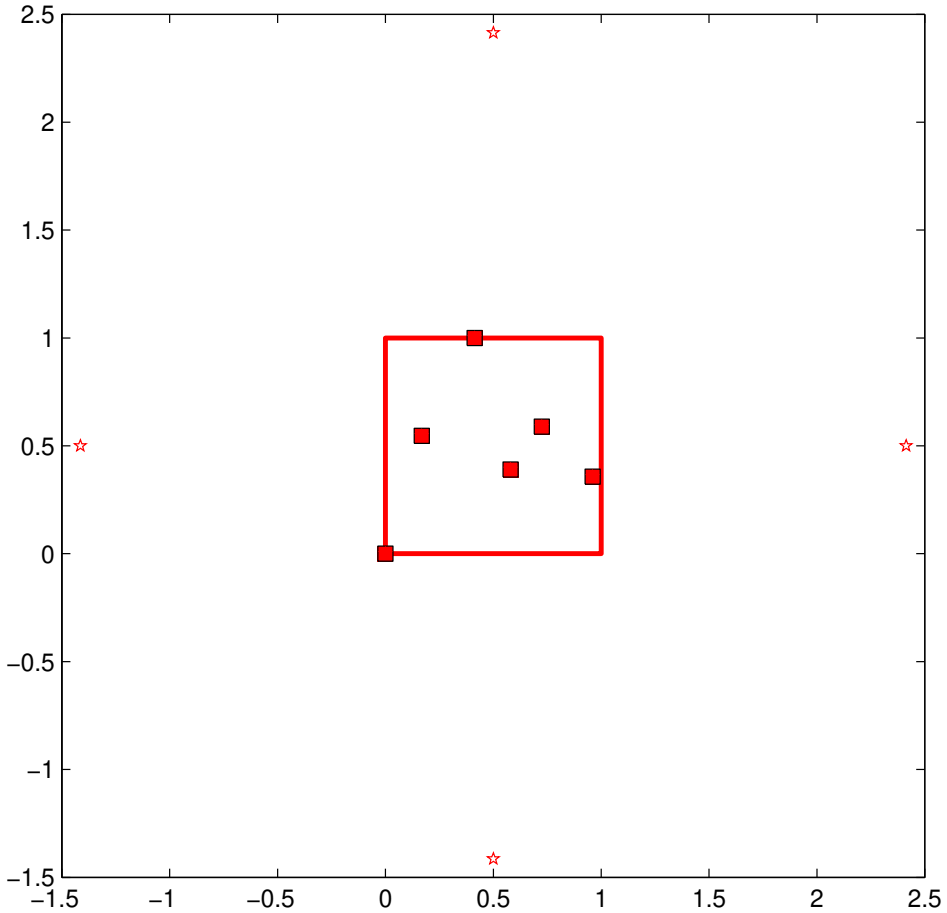
$\mathscr{X}$ = **polytope in** $\mathbb{R}^d$, see Cortés and Bullo (2005, 2009)

- Partition $\mathbb{R}^d$ into n cells $\mathcal{C}_i$ containing points closest to $\mathbf{x}_i$ than to any other site in X_n
- Each $\mathcal{C}_i$ = convex polyhedron in $\mathbb{R}^d$ (some are open and infinite)
- $\mathscr{X}$ is a polytope of $\mathbb{R}^d \Rightarrow \mathcal{C}_i \cap \mathscr{X}$ = polytope $\rightarrow$ tessellation of $\mathscr{X}$ into n bounded convex polyhedra
- $\max_{\mathbf{x}\in\mathscr{X}} d(\mathbf{x}, X_n)$ is attained when $\mathbf{x}$ is a vertex of one of these polyhedra
- Take $\mathscr{X}_Q$ = collection of these vertices
- $Q = \mathcal{O}(n^{\lceil d/2 \rceil}) \rightarrow$ small d only
- Avoid infinite cells by adding a few (at least $d+1$) generators $\mathbf{x}'_j$ out of $\mathscr{X}$, far enough from $\mathscr{X}$ to ensure that the corresponding cells do not intersect $\mathscr{X}$

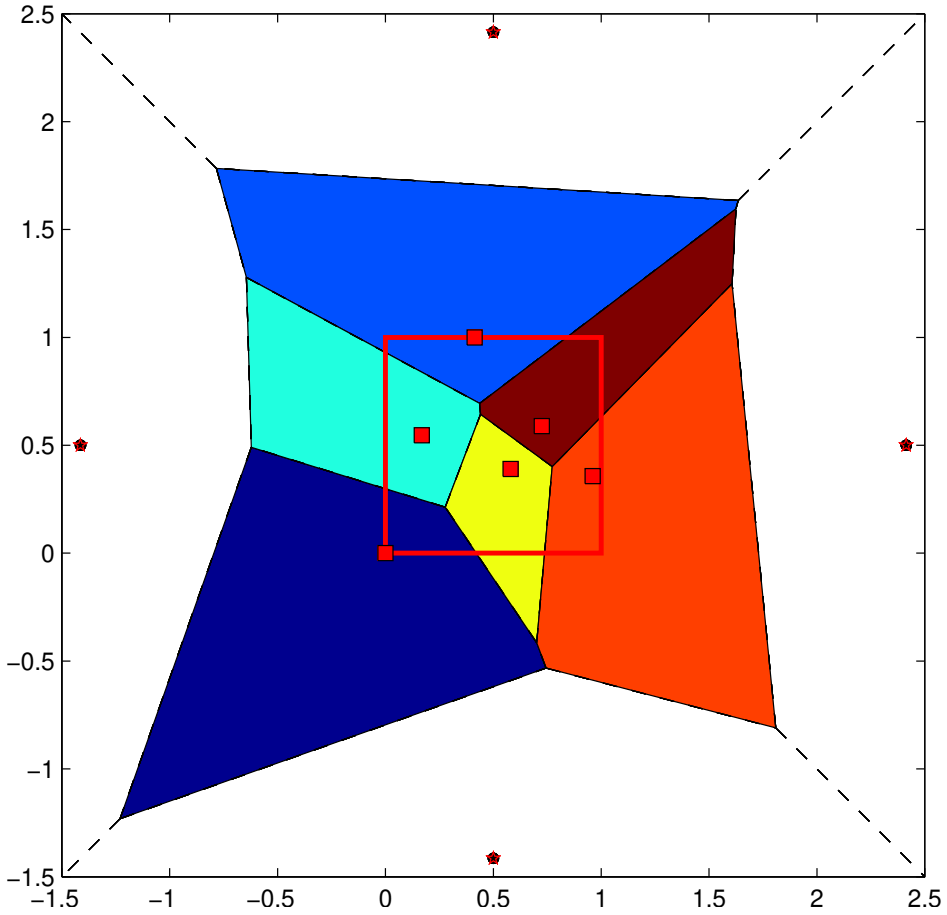
$\mathscr{X}_Q = \{$ vertices of Voronoï cells truncated to $\mathscr{X}$ $\}$



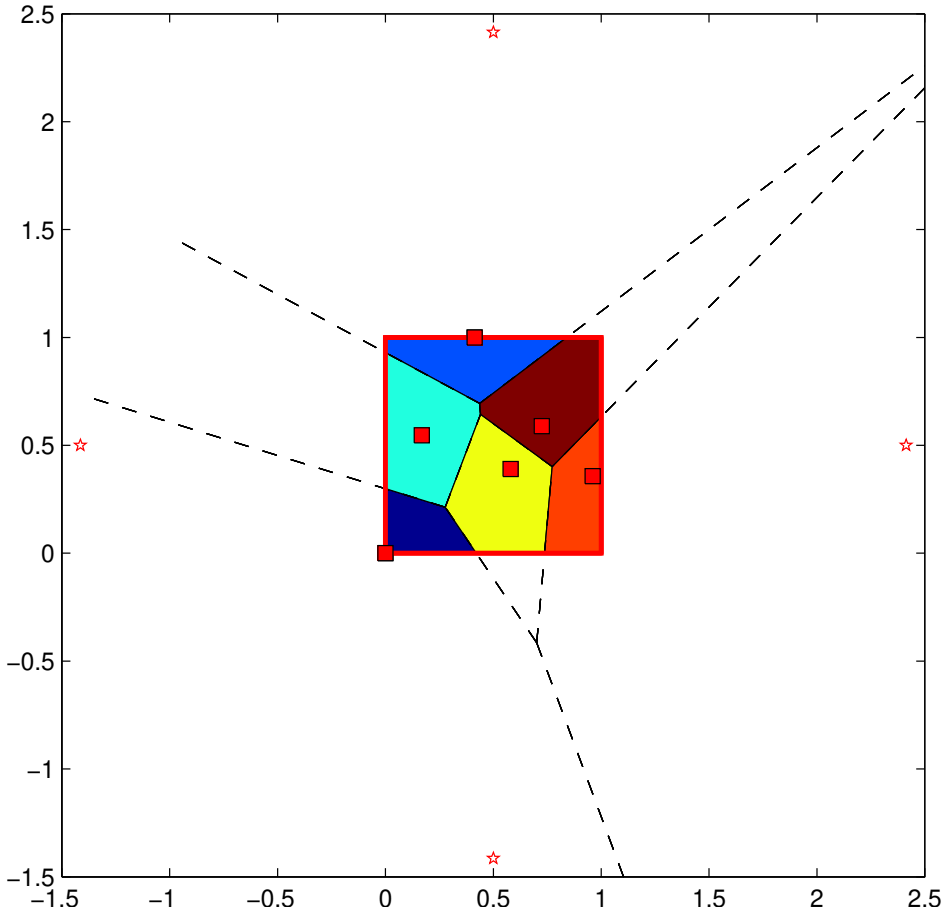
$\mathscr{X}_Q = \{$ vertices of Voronoï cells truncated to $\mathscr{X}$ $\}$



$\mathscr{X}_Q = \{$ vertices of Voronoï cells truncated to $\mathscr{X}$ $\}$



$\mathscr{X}_Q = \{$ vertices of Voronoï cells truncated to $\mathscr{X}$ $\}$



$n = 6$ points, 6 cells, $Q = 14$ vertices $\mathbf{x}^{(k)}$ tested for $\min_i \|\mathbf{x}^{(k)} - \mathbf{x}_i\|$

C/ Estimation via MCMC

2 ideas: extreme-value theory + multilevel splitting

C.a) Borrow results from extreme-value theory used in global optimization (Zhigljavsky and Žilinskas, 2007, Chap. 2), (Zhigljavsky and Hamilton, 2010)

- Q points $\mathbf{x}^{(j)}$ i.i.d. in $\mathscr{X}$, compute the Q distances $d_j = d(\mathbf{x}^{(j)}, X_n)$, associated order statistics $d_{1:Q} \geq d_{2:Q} \geq \cdots \geq d_{Q:Q}$

C/ Estimation via MCMC

2 ideas: extreme-value theory + multilevel splitting

C.a) Borrow results from extreme-value theory used in global optimization (Zhigljavsky and Žilinskas, 2007, Chap. 2), (Zhigljavsky and Hamilton, 2010)

- Q points $\mathbf{x}^{(j)}$ i.i.d. in $\mathscr{X}$, compute the Q distances $d_j = d(\mathbf{x}^{(j)}, X_n)$, associated order statistics $d_{1:Q} \geq d_{2:Q} \geq \cdots \geq d_{Q:Q}$
- k fixed, $1 \leq k \leq Q$ (e.g., $k = \max\{10, d\}$, $Q \gg d$), estimate $\Phi_{mM}(X_n)$ by

$$\widehat{\Phi}_{mM}(X_n) = d_{1:Q} + C_k(d_{1:Q} - d_{k:Q})$$

where $C_k = b_1/(b_k - b_1)$ with $b_i = \Gamma(i + 1/d)/\Gamma(i)$.
Also, the asymptotic confidence level of

$$I_{k,\delta} = \left[d_{1:Q}, d_{1:Q} + \frac{d_{1:Q} - d_{k:Q}}{(1 - \delta^{1/k})^{-1/d} - 1} \right]$$

tends to $1 - \delta$ for $Q \to \infty$

C/ Estimation via MCMC

2 ideas: extreme-value theory + multilevel splitting

C.a) Borrow results from extreme-value theory used in global optimization (Zhigljavsky and Žilinskas, 2007, Chap. 2), (Zhigljavsky and Hamilton, 2010)

- Q points $\mathbf{x}^{(j)}$ i.i.d. in $\mathscr{X}$, compute the Q distances $d_j = d(\mathbf{x}^{(j)}, X_n)$, associated order statistics $d_{1:Q} \geq d_{2:Q} \geq \cdots \geq d_{Q:Q}$
- k fixed, $1 \leq k \leq Q$ (e.g., $k = \max\{10, d\}$, $Q \gg d$), estimate $\Phi_{mM}(X_n)$ by

$$\widehat{\Phi}_{mM}(X_n) = d_{1:Q} + C_k(d_{1:Q} - d_{k:Q})$$

where $C_k = b_1/(b_k - b_1)$ with $b_i = \Gamma(i + 1/d)/\Gamma(i)$.
Also, the asymptotic confidence level of

$$I_{k,\delta} = \left[d_{1:Q}, d_{1:Q} + \frac{d_{1:Q} - d_{k:Q}}{(1 - \delta^{1/k})^{-1/d} - 1} \right]$$

tends to $1 - \delta$ for $Q \to \infty$

- Precise estimation only for very large Q ➠ 2nd idea

C.b) the order statistics $d_{j:Q}$ for large j (small $d_{j:Q}$) are useless

➠ multilevel splitting algorithm

C.b) the order statistics $d_{j:Q}$ for large j (small $d_{j:Q}$) are useless

➠ multilevel splitting algorithm

- Replace all $\mathbf{x}^{(j)}$ at distance d_j from X_n less than some L_ℓ by points sampled independently (and uniformly) in the set $\mathscr{X}(L_\ell) = \{\mathbf{x} \in \mathscr{X} : d(\mathbf{x}, X_n) > L_\ell\}$, for an increasing sequence of levels L_ℓ

C.b) the order statistics $d_{j:Q}$ for large j (small $d_{j:Q}$) are useless

➠ multilevel splitting algorithm

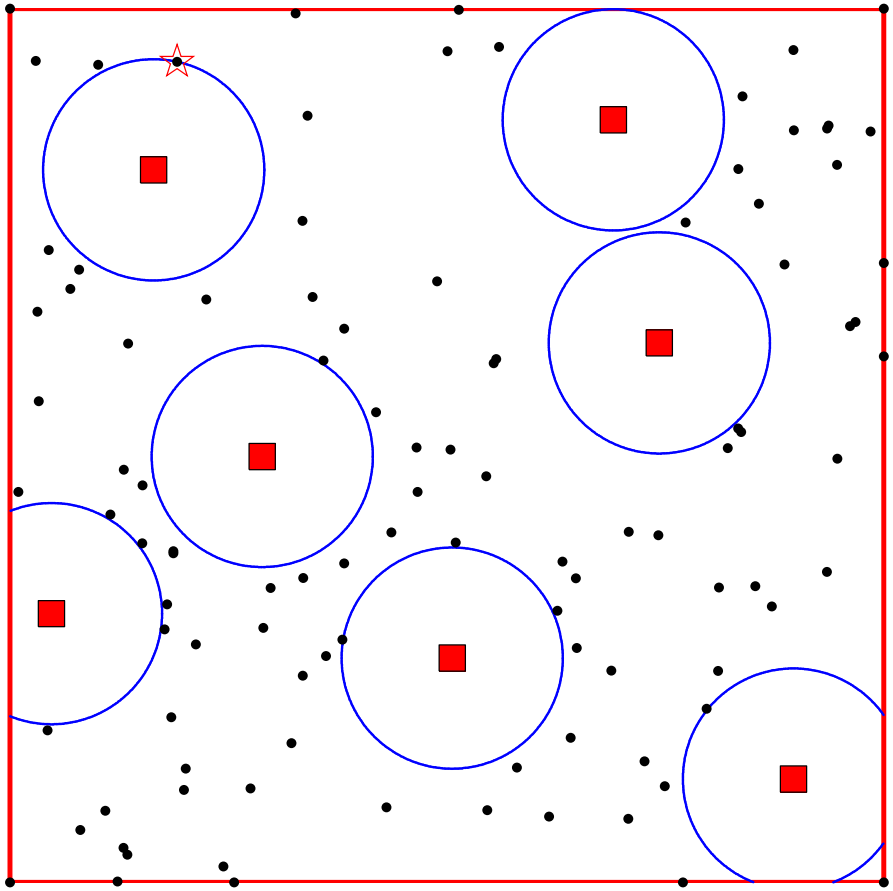
- Replace all $\mathbf{x}^{(j)}$ at distance d_j from X_n less than some L_ℓ by points sampled independently (and uniformly) in the set $\mathscr{X}(L_\ell) = \{\mathbf{x} \in \mathscr{X} : d(\mathbf{x}, X_n) > L_\ell\}$, for an increasing sequence of levels L_ℓ
- Choose the level sequence of Guyader et al. (2011): at step ℓ, the next level is $L_{\ell+1} = \min_{j=1,\ldots,Q} d_j$
  $\mathbf{x}_{j^*}$ (unique with probability one) such that $d_{j^*} = L_{\ell+1}$ is replaced by a new point sampled in $\mathscr{X}(L_{\ell+1})$
- Stop when $|I_{k,\delta}| < \epsilon \ll 1$ ($\delta = 0.05$, say)

C.b) the order statistics $d_{j:Q}$ for large j (small $d_{j:Q}$) are useless

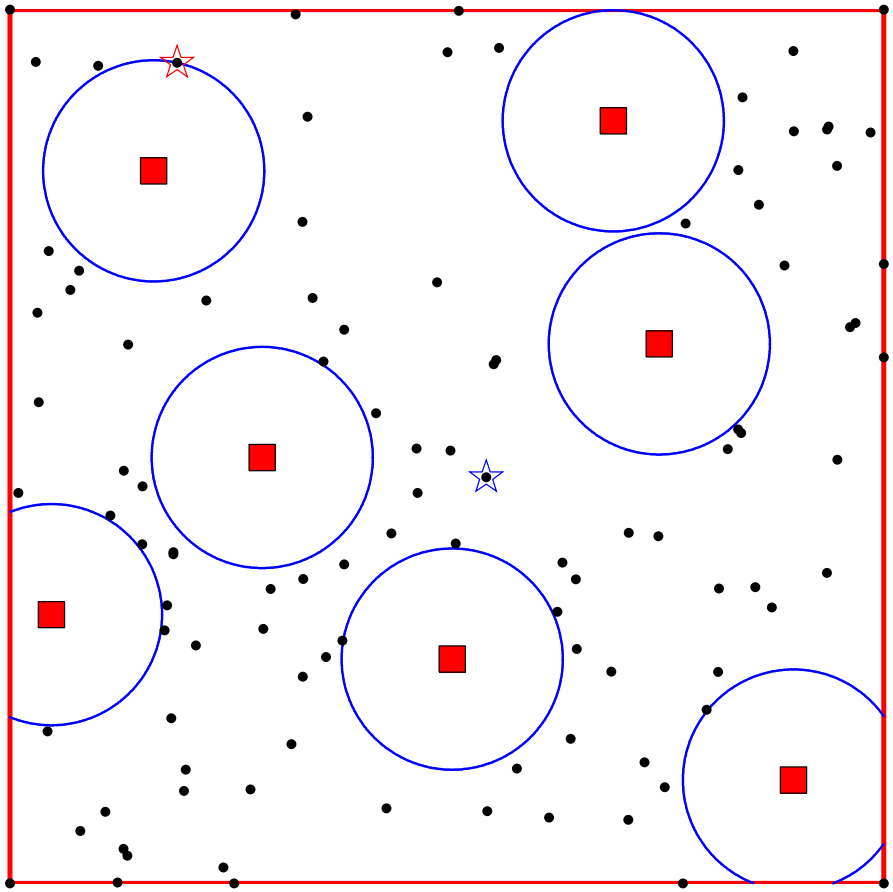
➠ multilevel splitting algorithm

- Replace all $\mathbf{x}^{(j)}$ at distance d_j from X_n less than some L_ℓ by points sampled independently (and uniformly) in the set $\mathscr{X}(L_\ell) = \{\mathbf{x} \in \mathscr{X} : d(\mathbf{x}, X_n) > L_\ell\}$, for an increasing sequence of levels L_ℓ
- Choose the level sequence of Guyader et al. (2011): at step ℓ, the next level is $L_{\ell+1} = \min_{j=1,\ldots,Q} d_j$
  $\mathbf{x}_{j^*}$ (unique with probability one) such that $d_{j^*} = L_{\ell+1}$ is replaced by a new point sampled in $\mathscr{X}(L_{\ell+1})$
- Stop when $|I_{k,\delta}| < \epsilon \ll 1$ ($\delta = 0.05$, say)
- Sampling ("uniformly") in $\mathscr{X}(L)$ is difficult when L is large: use a MCMC method with Metropolis-Hastings transitions as in (Guyader et al., 2011):
  - first replace $\mathbf{x}_{j^*}$ by a $\mathbf{x}_{j^{**}}$ chosen at random among the other $\mathbf{x}_j$
  - second, perform K successive steps of a random walk $\mathbf{x} \to \text{Proj}_{\mathscr{X}}(\mathbf{x} + \mathbf{z})$, with $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \sigma \mathbf{I}_d)$, accept transition if and only if $d(\mathbf{x} + \mathbf{z}, X_n) > L_{\ell+1} = d_{j^*}$

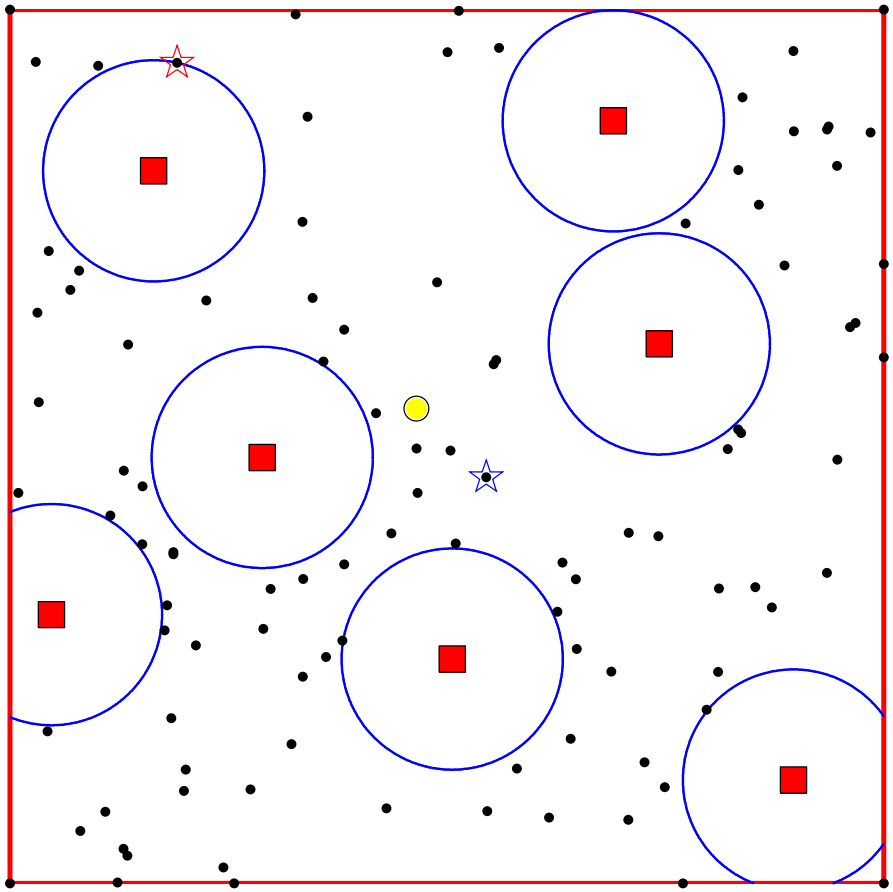
$\mathbf{x}_{j*}$ such that $d_{j*} = \min_j d_j$



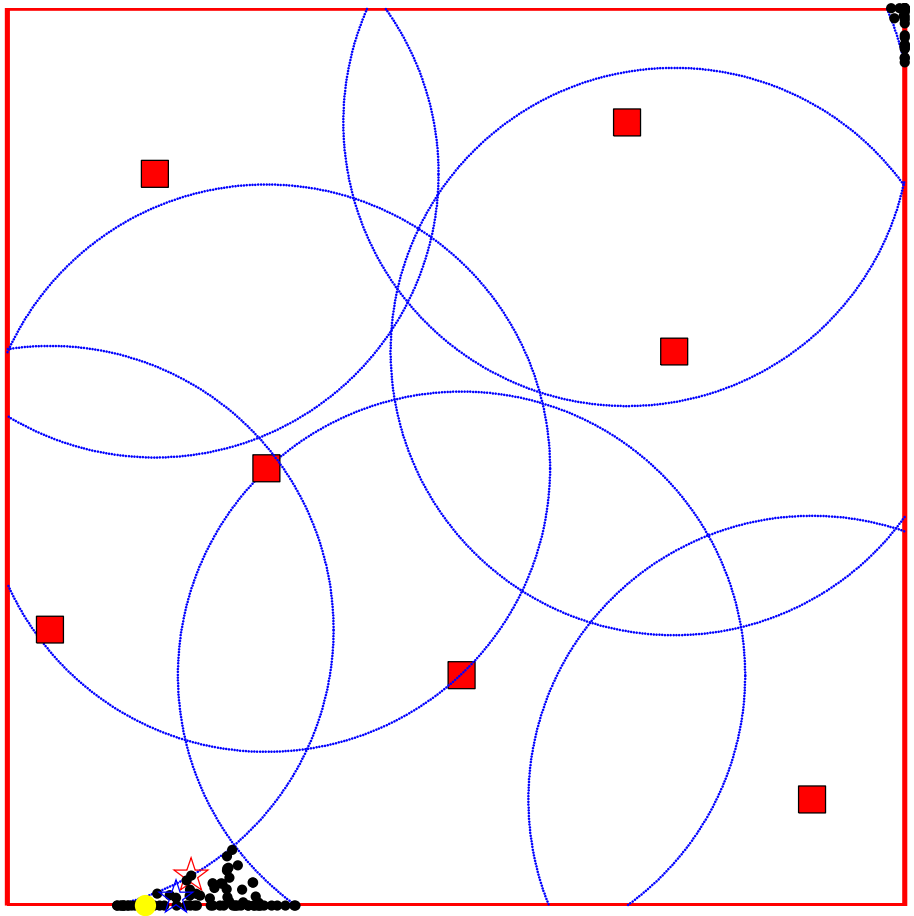
replace by $\mathbf{x}_{j^{**}}$ chosen at random among other $\mathbf{x}_j$



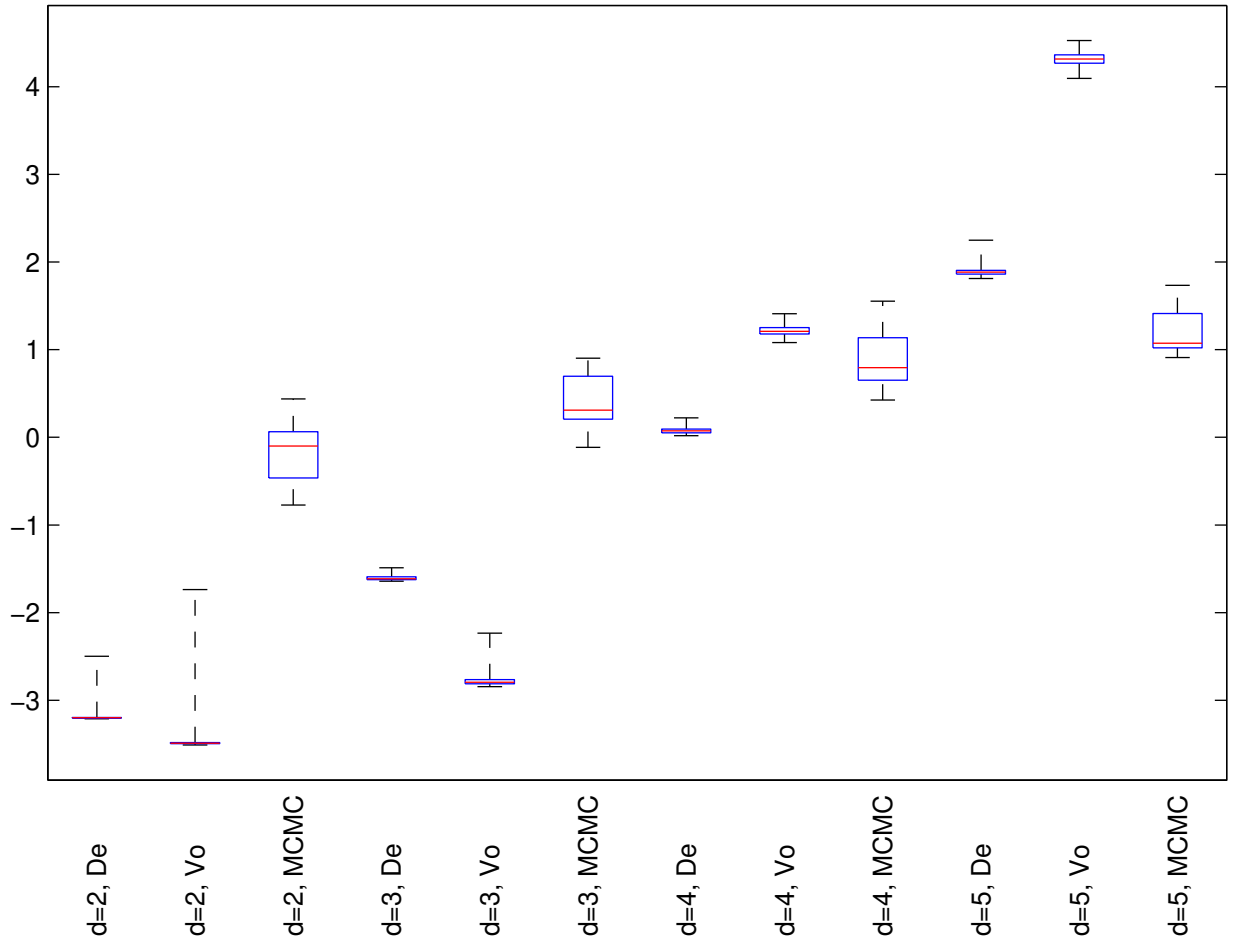
perform K successive steps of random walk



. . . after enough iterations



log(computing time)
$\mathscr{X} = [0,1]^d$, $n = 50$ ($\delta = 0.05$, $\epsilon = 0.001$, $K = 10$, $Q = nd$ for MCMC)



Minimization of $\Phi_{mM}(\mathbf{X}_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$?

(not convex, non differentiable)

Minimization of $\Phi_{mM}(\mathbf{X}_n) = \max_{\mathbf{x}\in\mathscr{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$?

(not convex, non differentiable)

- 0/ General global optimization method (e.g., simulated annealing): not promising
- A/ Voronoï tessellation + generalized gradient
- B/ `k-means` and centroids
- C/ Stochastic gradient

A/ d (very) small: Voronoï tessellation + generalized gradient
($\Phi_{mM}(\cdot)$ not differentiable, but Lipschitz, with constant 1)

$$X_n^{(k+1)} = X_n^{(k)} - \gamma_k \, \tilde{\nabla}_{\Phi_{mM}}(X_n^{(k)})$$

- $\gamma_k > 0$, $\lim_{k\to\infty} \gamma_k = 0$ and $\sum_k \gamma_k = \infty$
- all columns of $\tilde{\nabla}_{\Phi_{mM}}(X_n^{(k)})$ equal $\mathbf{0}$, except the i-th one equal to $(\mathbf{x}_i - \mathbf{x}^*)/\|\mathbf{x}_i - \mathbf{x}^*\|$, where $\|\mathbf{x}_i - \mathbf{x}^*\| = \Phi_{mM}(X_n)$
  ➠ move $\mathbf{x}_i$ towards $\mathbf{x}^*$

A/ d (very) small: Voronoï tessellation + generalized gradient
($\Phi_{mM}(\cdot)$ not differentiable, but Lipschitz, with constant 1)

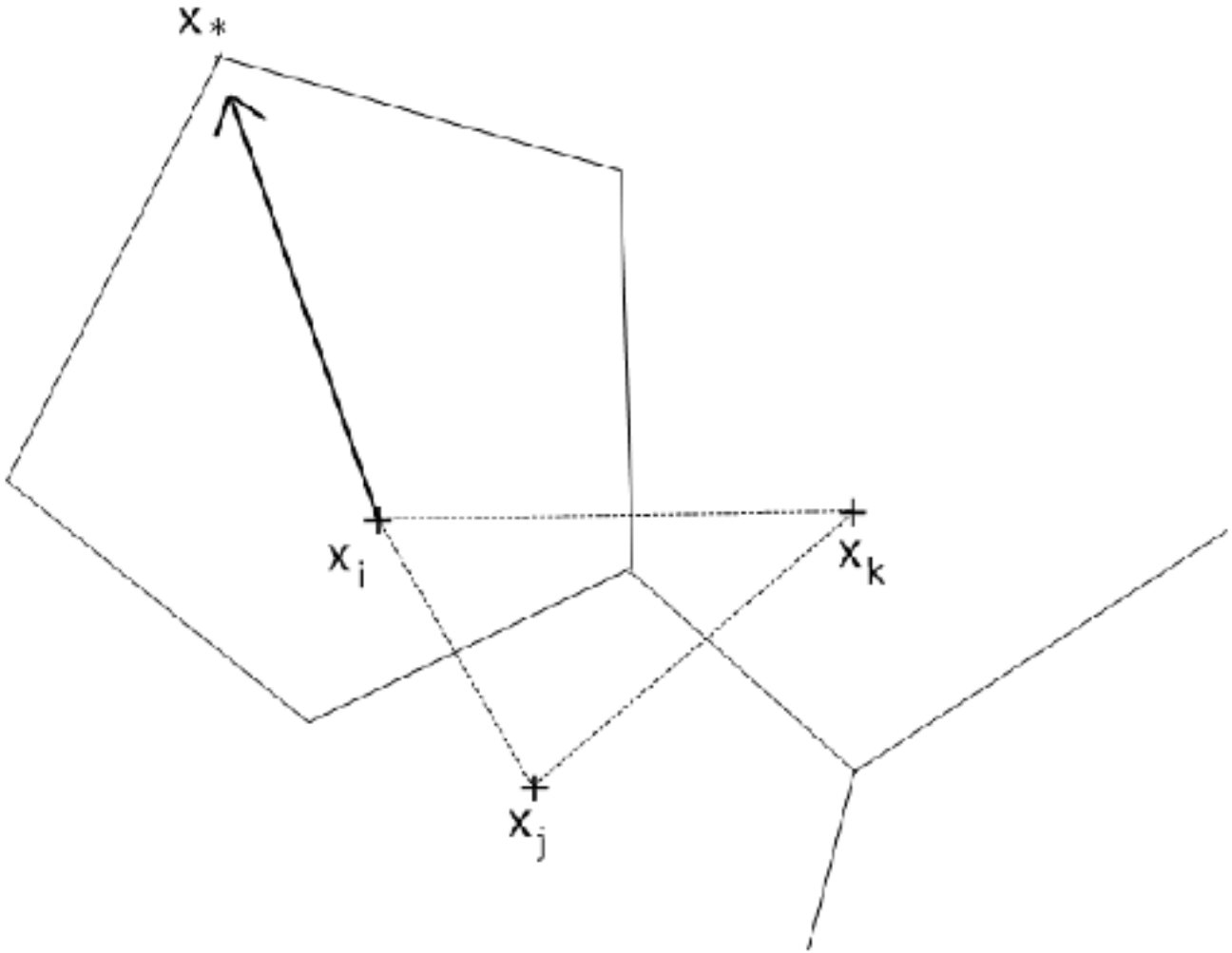
$$X_n^{(k+1)} = X_n^{(k)} - \gamma_k \, \tilde{\nabla}_{\Phi_{mM}}(X_n^{(k)})$$

- $\gamma_k > 0$, $\lim_{k\to\infty} \gamma_k = 0$ and $\sum_k \gamma_k = \infty$
- all columns of $\tilde{\nabla}_{\Phi_{mM}}(X_n^{(k)})$ equal $\mathbf{0}$, except the i-th one equal to $(\mathbf{x}_i - \mathbf{x}^*)/\|\mathbf{x}_i - \mathbf{x}^*\|$, where $\|\mathbf{x}_i - \mathbf{x}^*\| = \Phi_{mM}(X_n)$
  ➠ move $\mathbf{x}_i$ towards $\mathbf{x}^*$

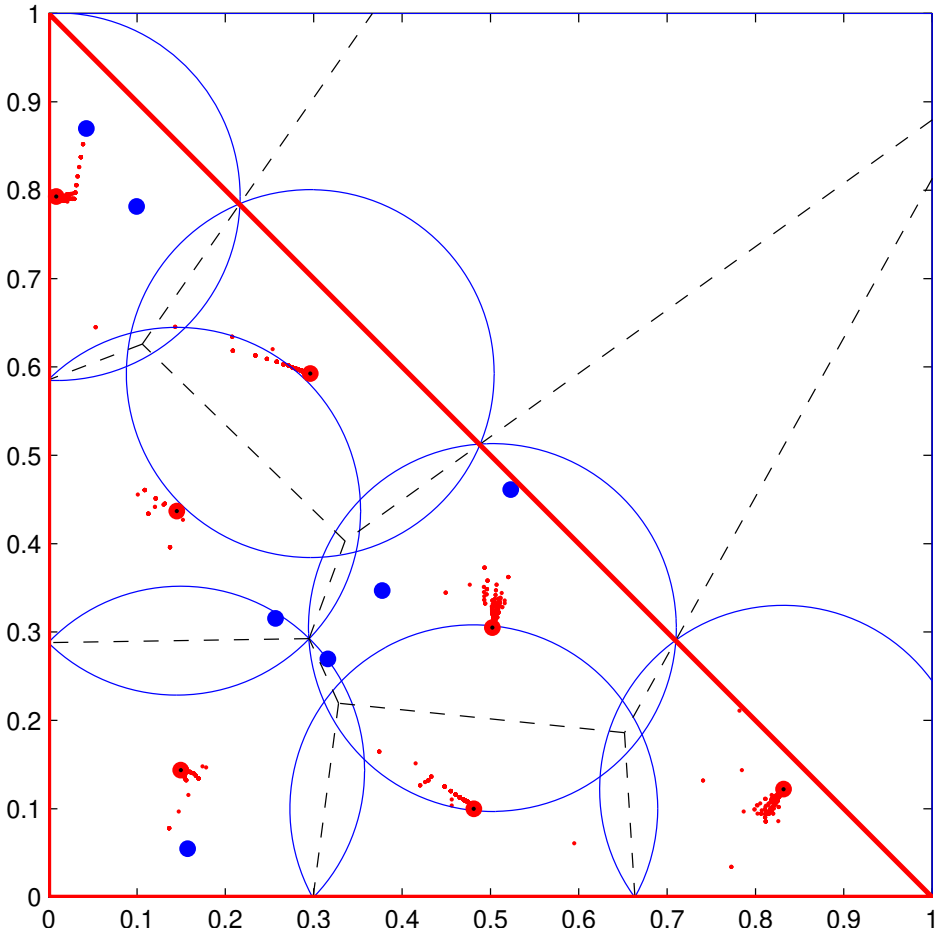
➠ one may also move each $\mathbf{x}_i$ towards the furthest point $\mathbf{x}^{*,i}$ in its Voronoï cell (Cortés and Bullo, 2005, 2009):

$$\mathbf{x}_i^{(k+1)} = \mathbf{x}_i^{(k)} - \gamma_{k,i}(\mathbf{x}_i^{(k)} - \mathbf{x}^{*,i})/\|\mathbf{x}_i^{(k)} - \mathbf{x}^{*,i}\|$$

Voronoï cells can be exact, or obtained by a discretization of $\mathscr{X}$ into a Q-point set $\mathscr{X}_Q$


x*
x_i
x_k
x_j

Ex: $\mathscr{X} =$ simplex $0 \leq x_1,\ 0 \leq x_2,\ x_1 + x_2 \leq 1,\ n = 7$ (radii $= \Phi_{mM}(\mathbf{X}_n)$)



B/ `k-means` and centroids

Minimize the L_2 energy functional

$$\mathcal{E}_2(\mathcal{T}_n, X_n) = \int_{\mathscr{X}} \left(\sum_{i=1}^{n} I_{\mathcal{C}_i}(\mathbf{x}) \, \|\mathbf{x} - \mathbf{x}_i\|^2 \right) \mathrm{d}\mathbf{x} = \sum_{i=1}^{n} \int_{\mathcal{C}_i} \|\mathbf{x} - \mathbf{x}_i\|^2 \, \mathrm{d}\mathbf{x}$$

where $\mathcal{T}_n = \{\mathcal{C}_i, \, i = 1, \ldots, n\}$ is a tessellation of $\mathscr{X}$

$I_{\mathcal{C}_i}$ = indicator function of $\mathcal{C}_i$

B/ `k-means` and centroids

Minimize the L_2 energy functional

$$\mathcal{E}_2(\mathcal{T}_n, X_n) = \int_{\mathscr{X}} \left(\sum_{i=1}^{n} I_{\mathcal{C}_i}(\mathbf{x}) \, \|\mathbf{x} - \mathbf{x}_i\|^2 \right) \mathrm{d}\mathbf{x} = \sum_{i=1}^{n} \int_{\mathcal{C}_i} \|\mathbf{x} - \mathbf{x}_i\|^2 \, \mathrm{d}\mathbf{x}$$

where $\mathcal{T}_n = \{\mathcal{C}_i, \, i = 1, \ldots, n\}$ is a tessellation of $\mathscr{X}$
$I_{\mathcal{C}_i}$ = indicator function of $\mathcal{C}_i$

Then (Du et al., 1999):

- $\mathcal{C}_i = \mathcal{V}(\mathbf{x}_i)$ = Voronoï region for the site $\mathbf{x}_i$, for all i
  $(\Rightarrow \mathcal{E}_2(\mathcal{T}_n, X_n) = \int_{\mathscr{X}} d^2(\mathbf{x}, X_n) \, \mathrm{d}\mathbf{x})$
- simultaneously $\mathbf{x}_i$ = centroid of $\mathcal{C}_i$ (center of gravity) for all i:
  $\mathbf{x}_i = (\int_{\mathcal{C}_i} \mathbf{x} \, \mathrm{d}\mathbf{x})/\mathrm{vol}(\mathcal{C}_i)$

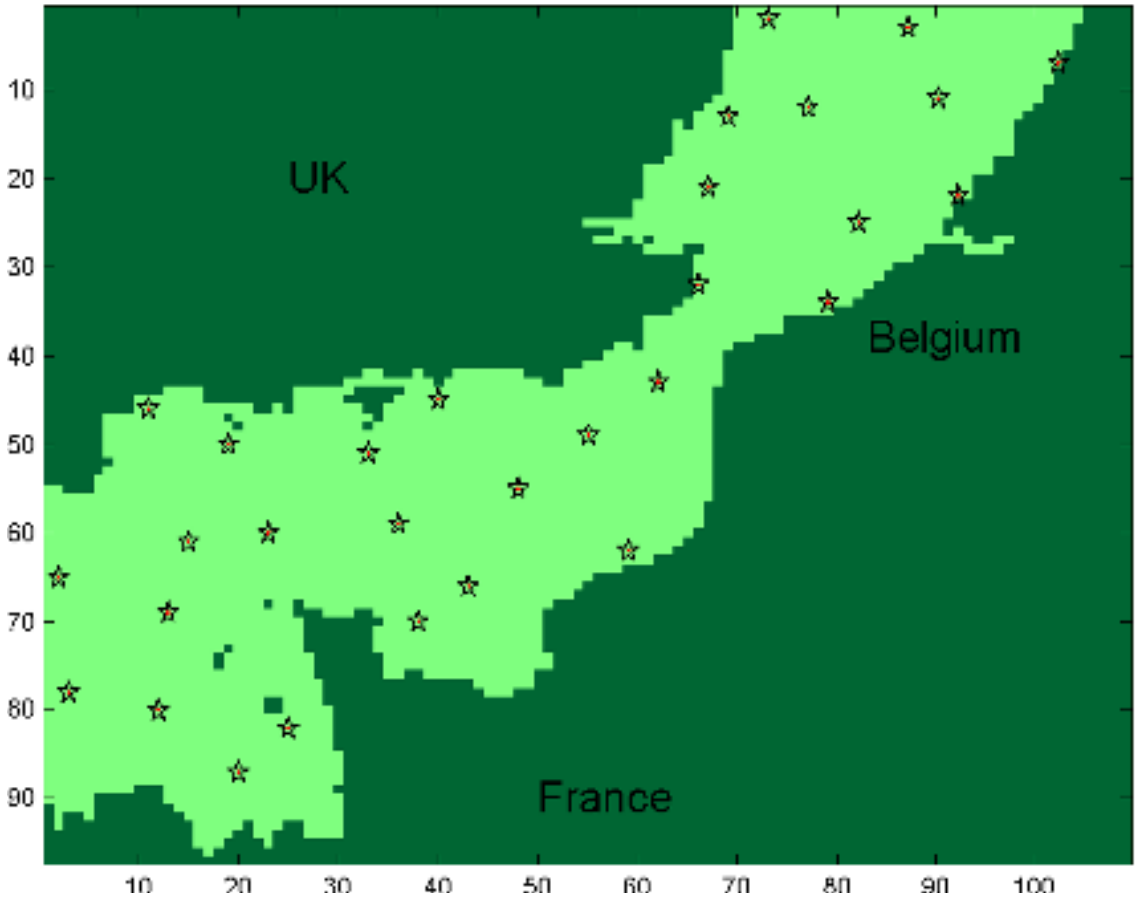
→ such a X_n should thus perform reasonably well in terms of space-filling (Lekivetz and Jones, 2015)

Lloyd's method (1982): (= fixed-point iterations)

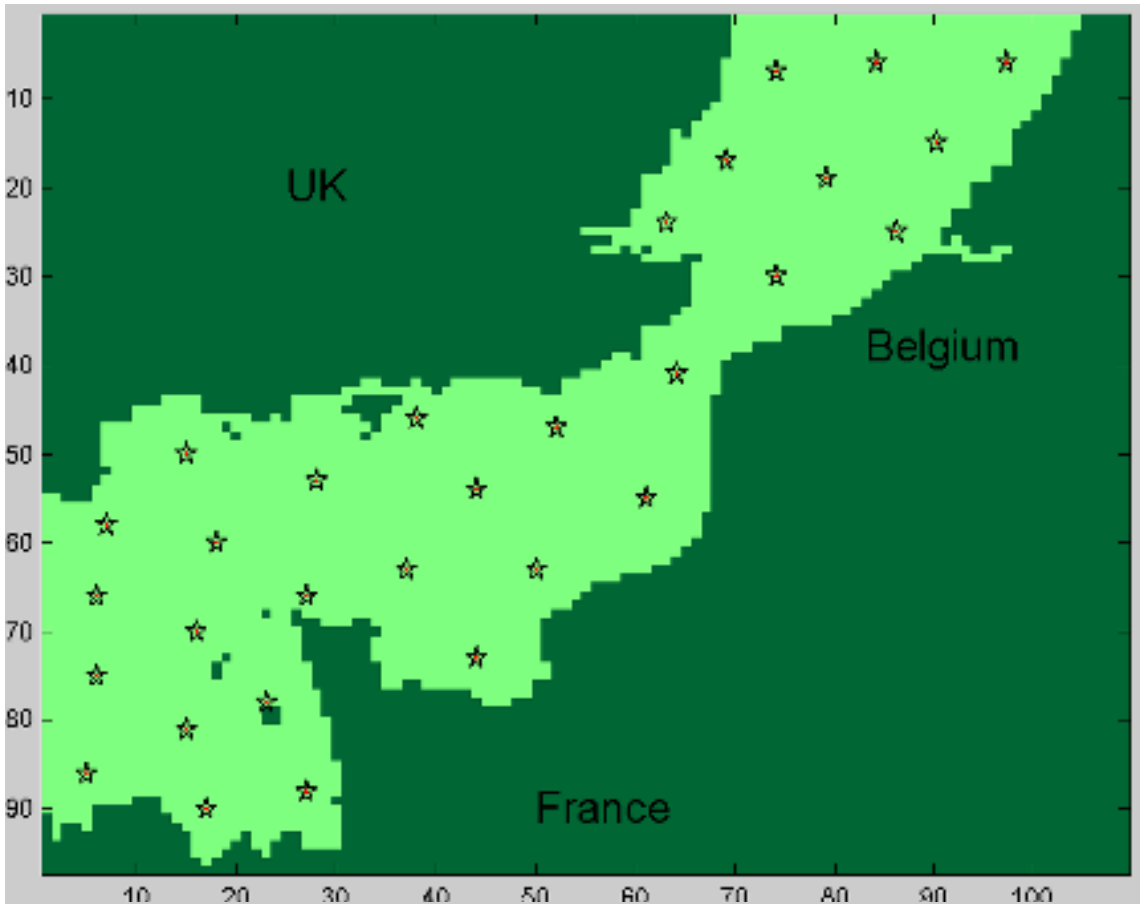
→ Move each $\mathbf{x}_i$ to the centroid of its own Voronoï cell, repeat ...

➠ Algorithmic geometry (Voronoï tessellation) if d very small,
use a finite set $\mathscr{X}_Q$ otherwise

30 points from Sobol' LDS



`k-means` clustering (30 clusters) of 1,000 point from Sobol' LDS



However... minimax-optimal design is related to the construction of a centroidal tessellation for

$$\mathcal{E}_q(\mathcal{T}_n, X_n) = \int_{\mathscr{X}} \left(\sum_{i=1}^{n} I_{\mathcal{C}_i}(\mathbf{x}) \, \|\mathbf{x} - \mathbf{x}_i\|^q \right) \mathrm{d}\mathbf{x} = \sum_{i=1}^{n} \int_{\mathcal{C}_i} \|\mathbf{x} - \mathbf{x}_i\|^q \, \mathrm{d}\mathbf{x}$$

for $q \to \infty$ (see (Mak and Joseph, 2016))

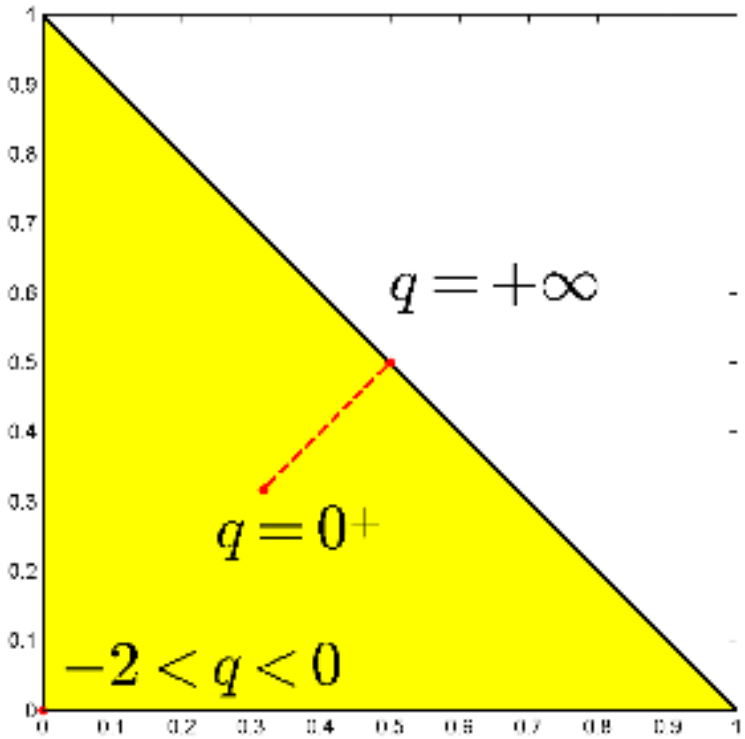
➠ use Chebyshev centers

However... minimax-optimal design is related to the construction of a centroidal tessellation for

$$\mathcal{E}_q(\mathcal{T}_n, X_n) = \int_{\mathcal{X}} \left(\sum_{i=1}^{n} l_{\mathcal{C}_i}(\mathbf{x}) \, \|\mathbf{x} - \mathbf{x}_i\|^q \right) \mathrm{d}\mathbf{x} = \sum_{i=1}^{n} \int_{\mathcal{C}_i} \|\mathbf{x} - \mathbf{x}_i\|^q \, \mathrm{d}\mathbf{x}$$

for $q \to \infty$ (see (Mak and Joseph, 2016))

➠ use Chebyshev centers

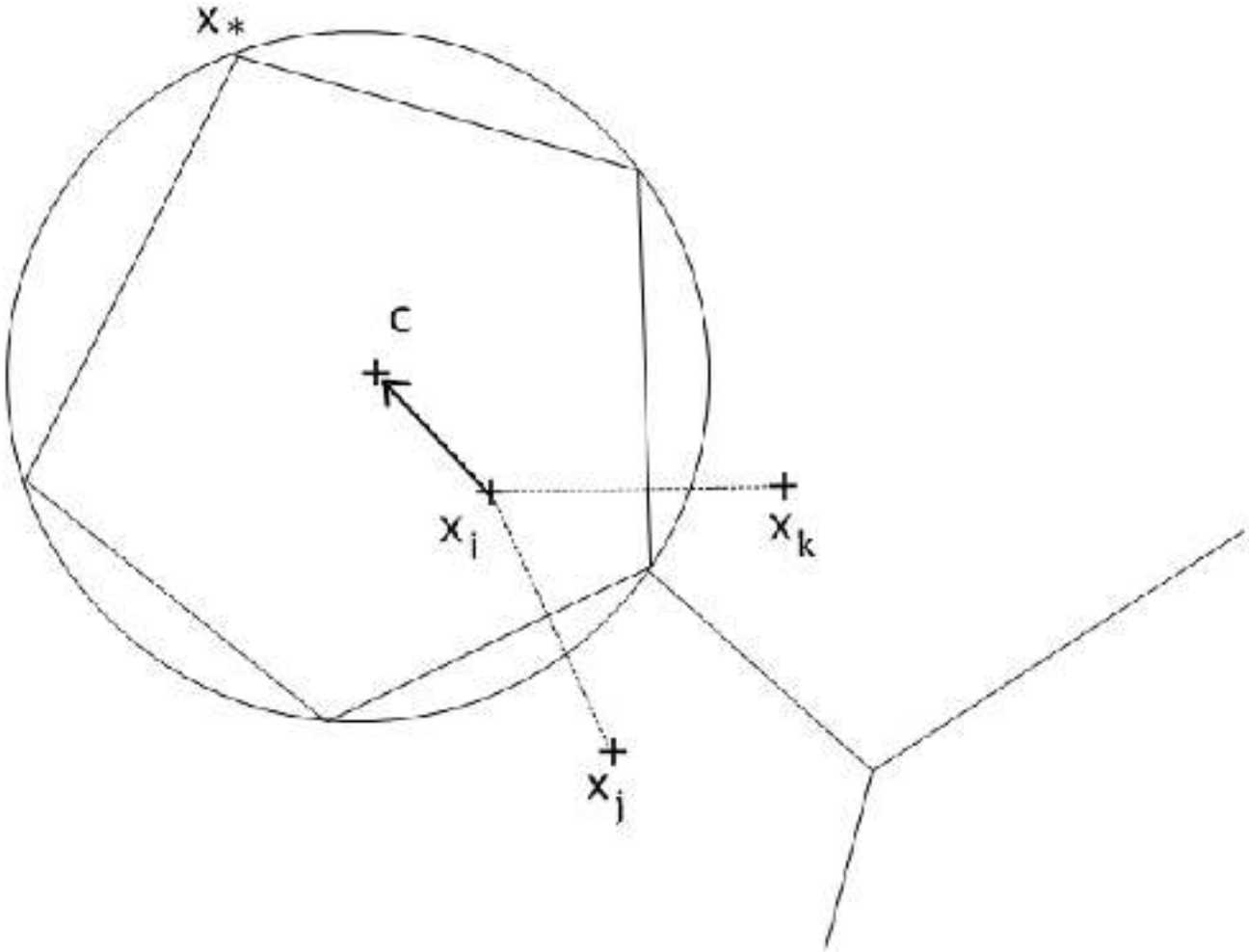


Variant of Lloyd's method:

0) Select $X_n^{(1)}$ and $\epsilon \ll 1$, set $k = 1$
1) Compute the Voronoï tessellation $\{\mathcal{V}_i,\ i = 1, \ldots, n\}$ of $\mathscr{X}$ (or $\mathscr{X}_Q$) based on $X_n^{(k)}$
2) For $i = 1, \ldots, n$
   ➤ determine the smallest ball $\mathscr{B}(\mathbf{c}_i, r_i)$ enclosing $\mathcal{V}_i$ (= convex QP problem)
   ➤ replace $\mathbf{x}_i$ by $\mathbf{c}_i$ in $X_n^{(k)}$ (Chebyshev center of $\mathcal{V}_i$)
3) if $\Phi_{mM}(\mathbf{X}_n^{(k)}) - \Phi_{mM}(\mathbf{X}_n^{(k+1)}) < \epsilon$, then stop; otherwise $k \leftarrow k + 1$, return to step 1

→ Move each $\mathbf{x}_i$ to the Chebyshev center of its own Voronoï cell, repeat ...

[$\Phi_{mM}(\mathbf{X}_n^{(k)})$ decreases monotonically, convergence to a local minimum (or a saddle point)]


x*
c
x_i
x_k
x_j

Determination of the smallest enclosing ball containing $\mathcal{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$
(vertices of a Voronoï cell, points of $\mathscr{X}_Q$ closest to $\mathbf{x}_i$):

$$\Leftrightarrow \text{ minimize } f(\mathbf{c}) = \max_{i=1,\ldots,N} \|\mathbf{z}_i - \mathbf{c}\|^2 \text{ with respect to } \mathbf{c} \in \mathbb{R}^d$$

Determination of the smallest enclosing ball containing $\mathcal{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$
(vertices of a Voronoï cell, points of $\mathscr{X}_Q$ closest to $\mathbf{x}_i$):

$$\Leftrightarrow \text{ minimize } f(\mathbf{c}) = \max_{i=1,\ldots,N} \|\mathbf{z}_i - \mathbf{c}\|^2 \text{ with respect to } \mathbf{c} \in \mathbb{R}^d$$

Direct problem = convex QP

Take any $\mathbf{c}_0 \in \mathbb{R}^d$, minimize $\|\mathbf{c} - \mathbf{c}_0\|^2 + t$

with respect to $(\mathbf{c}, t) \in \mathbb{R}^{d+1}$,

subject to $\|\mathbf{z}_i - \mathbf{c}_0\|^2 - 2(\mathbf{z}_i - \mathbf{c}_0)^\top(\mathbf{c} - \mathbf{c}_0) \leq t\,, \;\; i = 1, \ldots, N$

(N linear constraints)

Determination of the smallest enclosing ball containing $\mathcal{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$

Dual problem = similar to an optimal design problem:

maximize $\mathrm{trace}[\mathbf{V}(\xi)]$, with ξ a prob. measure on $\mathcal{Z}$,

$\mathbf{V}(\xi)$ = covariance matrix for ξ

center of the ball = $\mathbf{c}(\xi) = \int_{\mathcal{Z}} \mathbf{z}\, \xi(\mathrm{d}\mathbf{z})$

Determination of the smallest enclosing ball containing $\mathcal{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$

Dual problem = similar to an optimal design problem:

maximize $\mathrm{trace}[\mathbf{V}(\xi)]$, with ξ a prob. measure on $\mathcal{Z}$,

$\mathbf{V}(\xi)$ = covariance matrix for ξ

center of the ball = $\mathbf{c}(\xi) = \int_{\mathcal{Z}} \mathbf{z}\, \xi(\mathrm{d}\mathbf{z})$

→ Algorithms of the exchange-type (Yildirim, 2008)
(≈ Fedorov algorithm for D-optimal design: optimal step length is available)

→ One can remove inessential points from $\mathcal{Z}$: (P., 2017c)

➠ Combine this with the use of a standard QP solver for the direct problem

C/ Stochastic gradient (P., 2017)

d **is large**: Lloyd's algorithm cannot be used (computational geometry is too complicated, regular grids or LDS are not dense enough)

$$\text{minimize } \mathcal{E}_q{}^*(X_n) = \int_{\mathscr{X}} \left(\sum_{i=1}^{n} l_{\mathcal{V}_i}(\mathbf{x}) \, \|\mathbf{x} - \mathbf{x}_i\|^q \right) \mathrm{d}\mathbf{x}$$

with $\mathcal{V}_i$ = Voronoï region for the site $\mathbf{x}_i$

C/ Stochastic gradient (P., 2017)

d **is large**: Lloyd's algorithm cannot be used (computational geometry is too complicated, regular grids or LDS are not dense enough)

$$\text{minimize } \mathcal{E}_q{}^*(X_n) = \int_{\mathscr{X}} \left(\sum_{i=1}^{n} I_{\mathcal{V}_i}(\mathbf{x}) \, \|\mathbf{x} - \mathbf{x}_i\|^q \right) \mathrm{d}\mathbf{x}$$

with $\mathcal{V}_i$ = Voronoï region for the site $\mathbf{x}_i$

→ Stochastic gradient algorithm:
(MacQueen, 1967) for $q = 2$, (Cardot et al., 2012) for $q = 1$

0) $k = 1$, $X_n^{(1)}$, set $n_{i,0} = 0$ for all $i = 1, \ldots, n$

1) sample X uniformly distributed in $\mathscr{X}$

2) find $i^* = \arg\min_{i=1,\ldots,n} \|X - \mathbf{x}_i^{(k)}\|$, $n_{i^*,k} \leftarrow n_{i^*,k} + 1$ $\quad [\leftarrow X \in \text{cell } \mathcal{V}_i^*]$

3) $\mathbf{x}_{i^*}^{(k+1)} = \mathbf{x}_{i^*}^{(k)} - \gamma_{i^*,k} \underbrace{q\|X - \mathbf{x}_{i^*}^{(k)}\|^{q-2} (\mathbf{x}_{i^*}^{(k)} - X)}_{=\text{gradient}}$, $k \leftarrow k + 1$,

return to step 1, stop when $k = K$

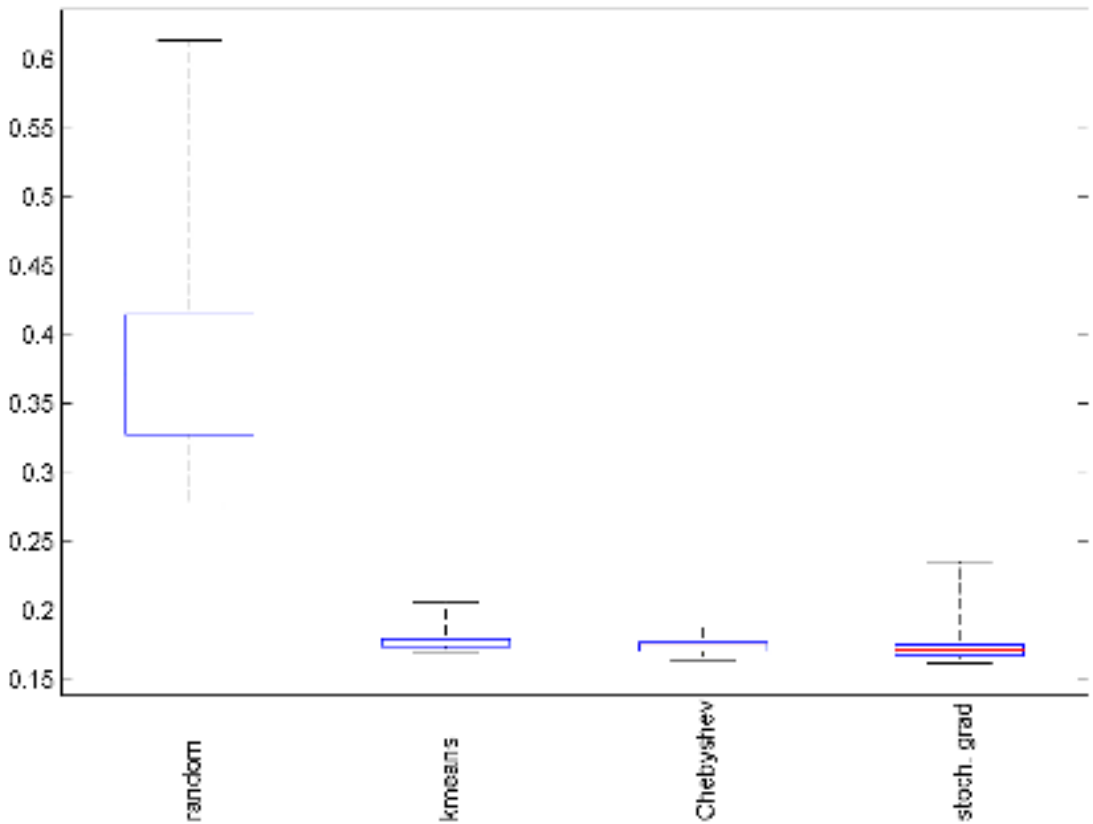
- Typical choice for $\gamma_{i^*,k} = c/n_{i^*,k}^{\alpha}$, with $\alpha \in (1/2, 1]$
  and consider $\widehat{X}_n = \frac{1}{K} \sum_{k=1}^{K} X_n^{(k)}$ when $\alpha < 1$
- Little information to store (no grid or other finite approximation of $\mathscr{X}$)
  $\rightarrow$ can also be used with large d

Example: $n = 10\,d$

all methods are initialized at the same random design, 100 repetitions

`k-means` and Lloyd's method with Chebyshev centers use 2^{d+8} points from a LDS (Sobol')

$$d = 2,\ n = 20$$

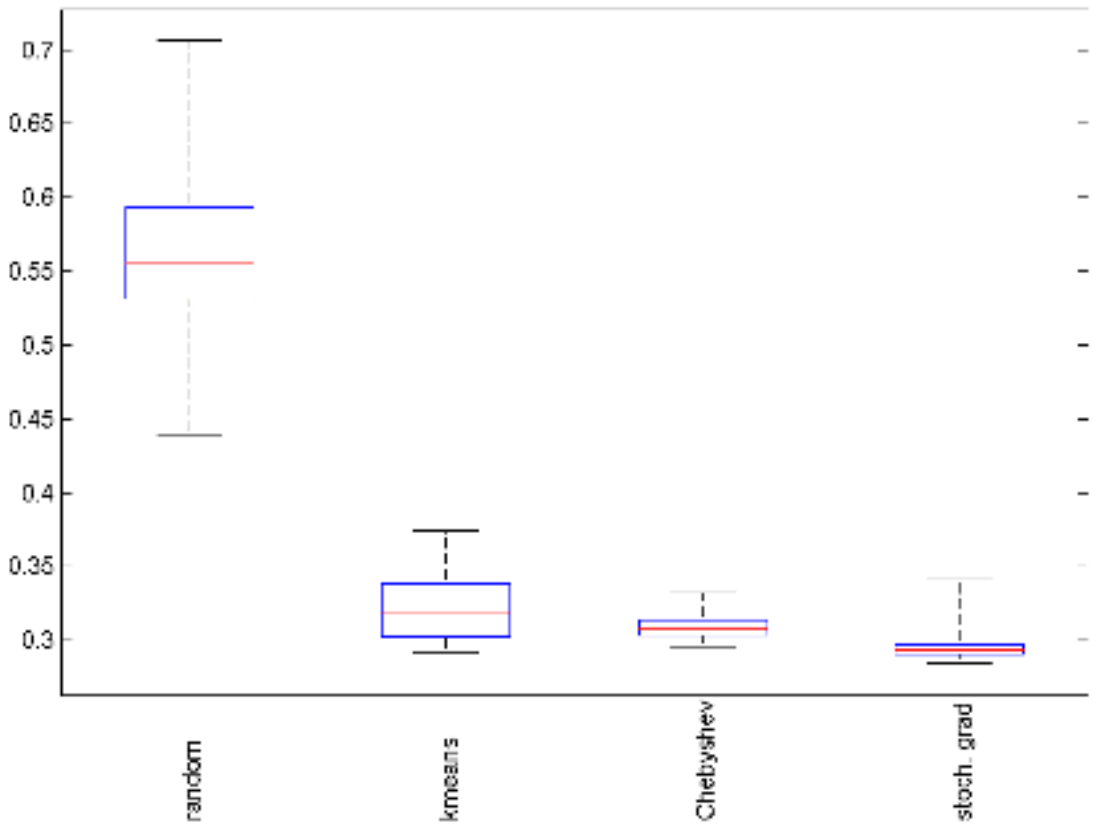


Example: $n = 10\,d$

all methods are initialized at the same random design, 100 repetitions

`k-means` and Lloyd's method with Chebyshev centers use 2^{d+8} points from a LDS (Sobol')

$$d = 3,\ n = 30$$

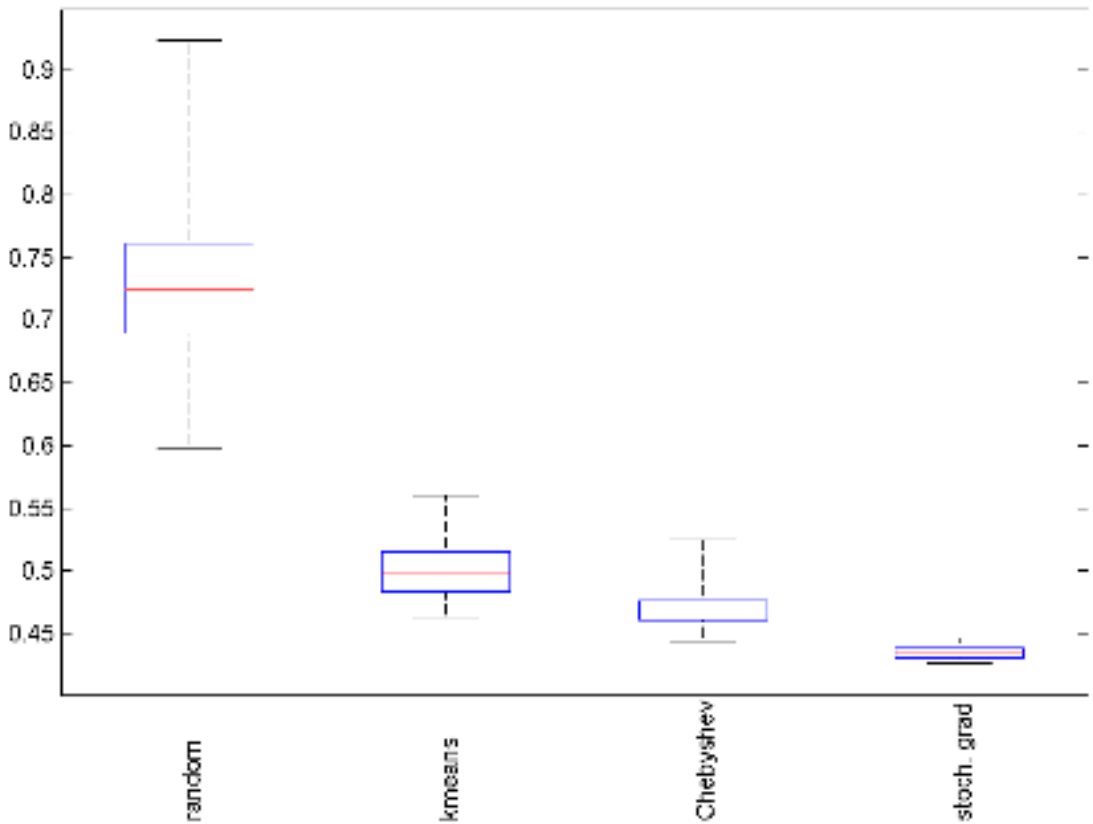


Example: $n = 10\,d$

all methods are initialized at the same random design, 100 repetitions

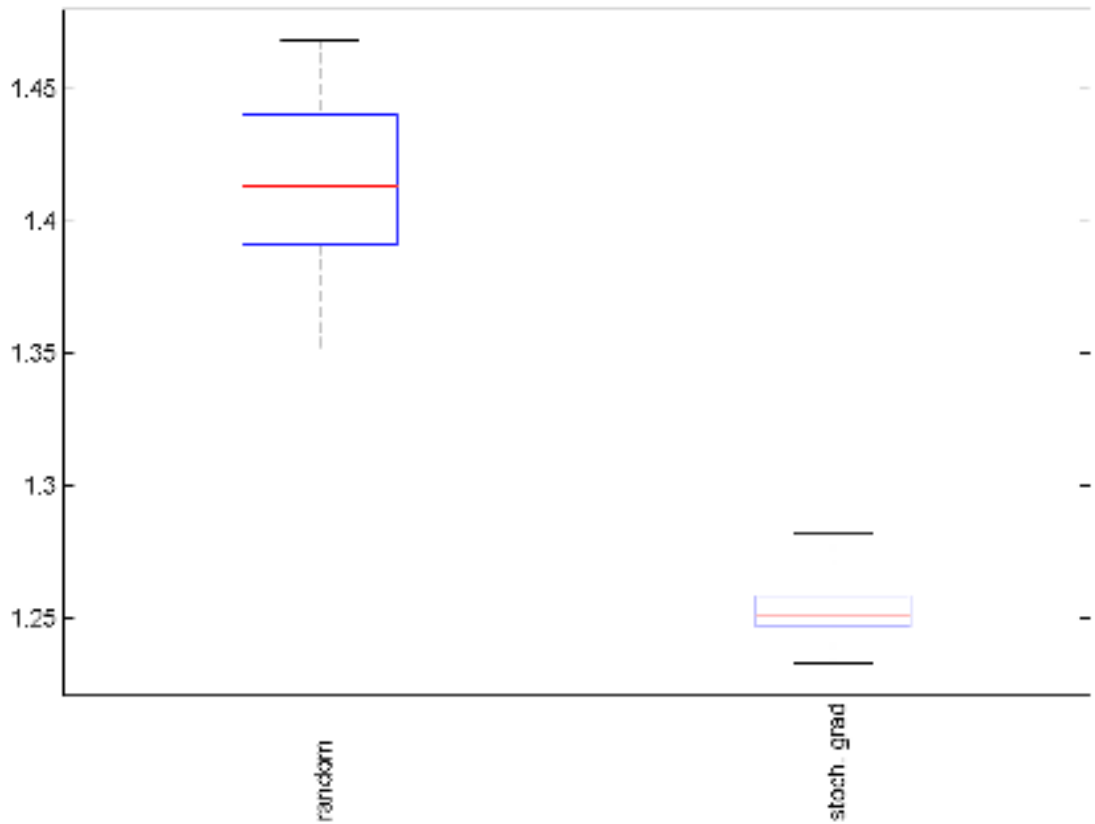
`k-means` and Lloyd's method with Chebyshev centers use 2^{d+8} points from a LDS (Sobol')

$$d = 4,\ n = 40$$



Example:

$$d = 10,\ n = 100$$



1.4 **Maximin criterion** $\Phi_{Mm}(\mathbf{X}_n) = \min_{i \neq j} d_{ij} = \min_{i \neq j} \|\mathbf{x}_i - \mathbf{x}_j\|$

Easy to compute (from distances d_{ij} between pairs of points)
$\Phi_{Mm}(\mathbf{X}_n)$ = minimum of convex functions $\Rightarrow$ not concave, non differentiable

- Use a global optimization method (e.g., simulated annealing)
- Local descent with some ad'hoc initialization
  (e.g., random multistart)

1.4 **Maximin criterion** $\Phi_{Mm}(\mathbf{X}_n) = \min_{i \neq j} d_{ij} = \min_{i \neq j} \|\mathbf{x}_i - \mathbf{x}_j\|$

Easy to compute (from distances d_{ij} between pairs of points)
$\Phi_{Mm}(\mathbf{X}_n)$ = minimum of convex functions $\Rightarrow$ not concave, non differentiable

- Use a global optimization method (e.g., simulated annealing)
- Local descent with some ad'hoc initialization
  (e.g., random ➠ multistart)

Difficult problem, but:

$\Phi_{Mm}(\cdot)$ is global Lipschitz (with constant $\sqrt{2}$)
$\rightarrow$ $\Phi_{Mm}(\cdot)$ is differentiable almost everywhere (Cortés and Bullo, 2005, 2009)

Sub-differential $\partial\Phi_{Mm}(\mathbf{X}_n)$ easy to compute:

$$\partial\Phi_{Mm}(\mathbf{X}_n) = co\{\partial\Phi_{Mm\ ij}(\mathbf{X}_n) : \|\mathbf{x}_i - \mathbf{x}_j\| = \min_{k\neq\ell} \|\mathbf{x}_k - \mathbf{x}_\ell\|\}$$

with $\Phi_{Mm\ ij}(\mathbf{X}_n) = \|\mathbf{x}_i - \mathbf{x}_j\|$ and

$$\text{sub-gradient } \partial\Phi_{Mm\ ij}(\mathbf{X}_n) = (0,\ldots,0,\underbrace{\frac{\mathbf{x}_i - \mathbf{x}_j}{\|\mathbf{x}_i - \mathbf{x}_j\|}}_{i\text{th position}},0,\ldots,0,\underbrace{-\frac{\mathbf{x}_i - \mathbf{x}_j}{\|\mathbf{x}_i - \mathbf{x}_j\|}}_{j\text{th position}},0,\ldots,0)$$

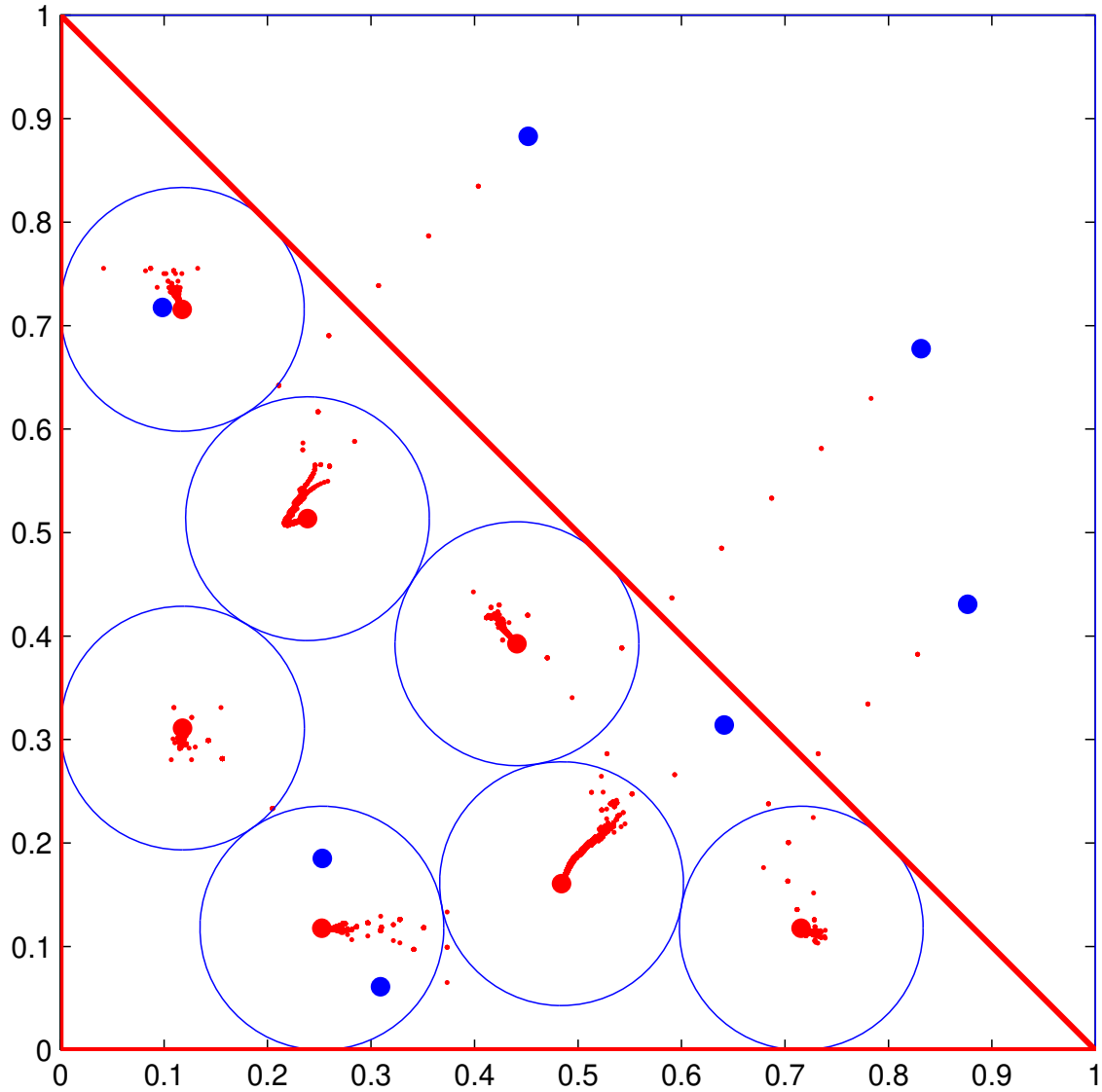
Sub-gradient algorithm to maximize $\Phi_{Mm}(\cdot)$:

$$\boxed{\mathbf{X}_n^{k+1} = \text{Proj}_{\mathscr{X}} \left[\mathbf{X}_n^k + \gamma_k \, \partial \Phi_{Mm \; ij}(\mathbf{X}_n^k) \right]}$$

for i, j such that $\|\mathbf{x}_i - \mathbf{x}_j\| = \Phi_{Mm}(\mathbf{X}_n^k)$
and $\gamma_k \searrow 0$, $\sum_k \gamma_k = \infty$

We can also force all points to remain far away from the boundary of $\mathscr{X}$:
$\Phi_{Mm \; B(\mathscr{X})}(\mathbf{X}_n) = \min\{\Phi_{Mm}(\mathbf{X}_n), 2 \min_i d[\mathbf{x}_i, \text{boundary}(\mathscr{X})]\}$

Ex: local maximization of $\Phi_{Mm\ B(\mathscr{X})}(\mathbf{X}_n)$, $n = 7$, $\mathscr{X}$ = simplex $0 \leq x_1$, $0 \leq x_2$, $x_1 + x_2 \leq 1$
(radius $= \frac{1}{2}\,\Phi_{Mm\ B(\mathscr{X})}(\mathbf{X}_n)$)



A few alternatives:

1) **Billiards** (Lubachevsky and Stillinger, 1990; Lubachevsky, 1991)

Principle :

$\mathbf{X}_n \to n$ balls in $\mathscr{X}$

random initial velocities

elastic collisions between balls and against boundaries

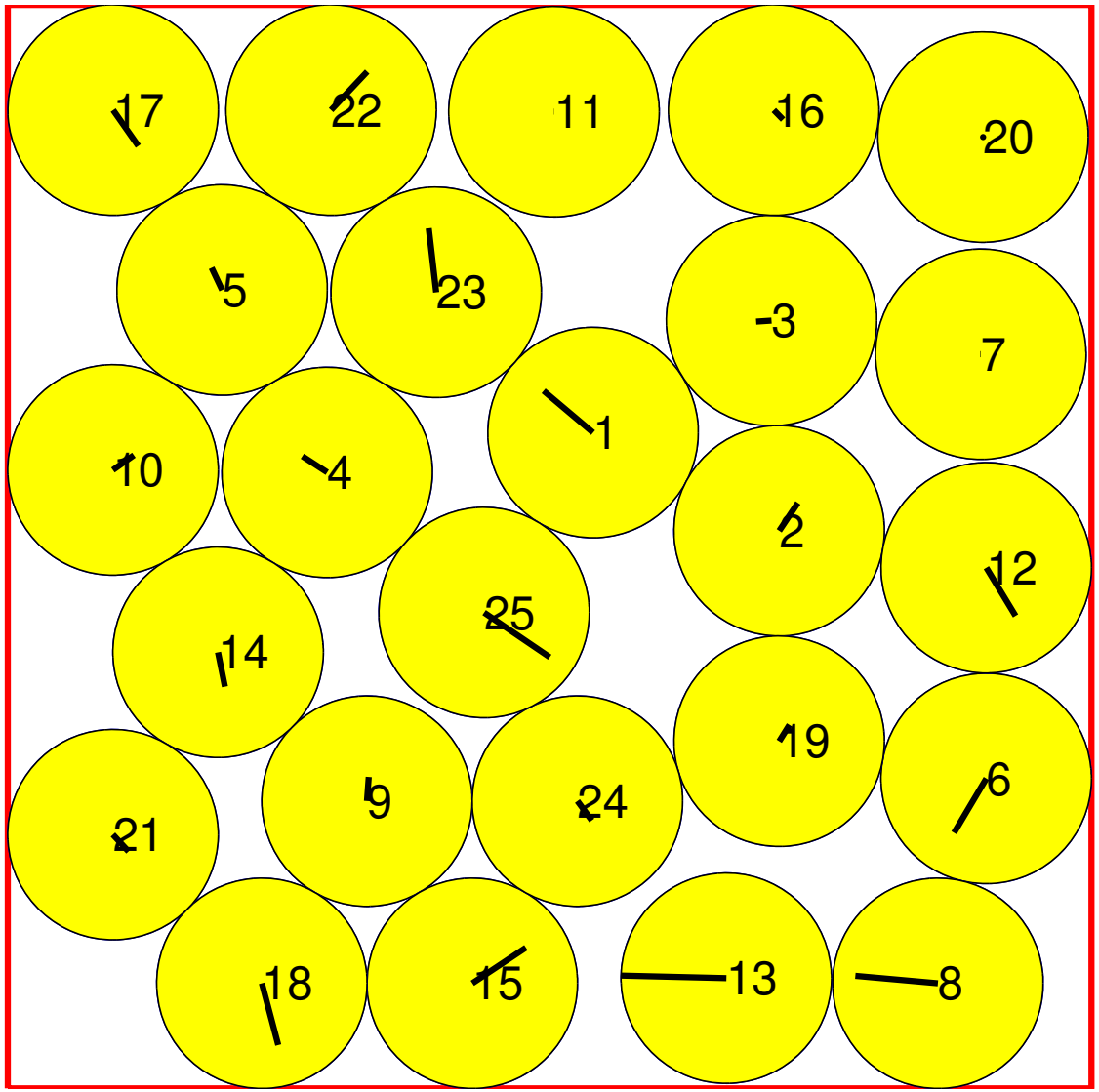
balls radius $R(t)$ linearly increasing with time t

➠ jamming occurs for a local max. of $\Phi_{Mm}(\cdot)$

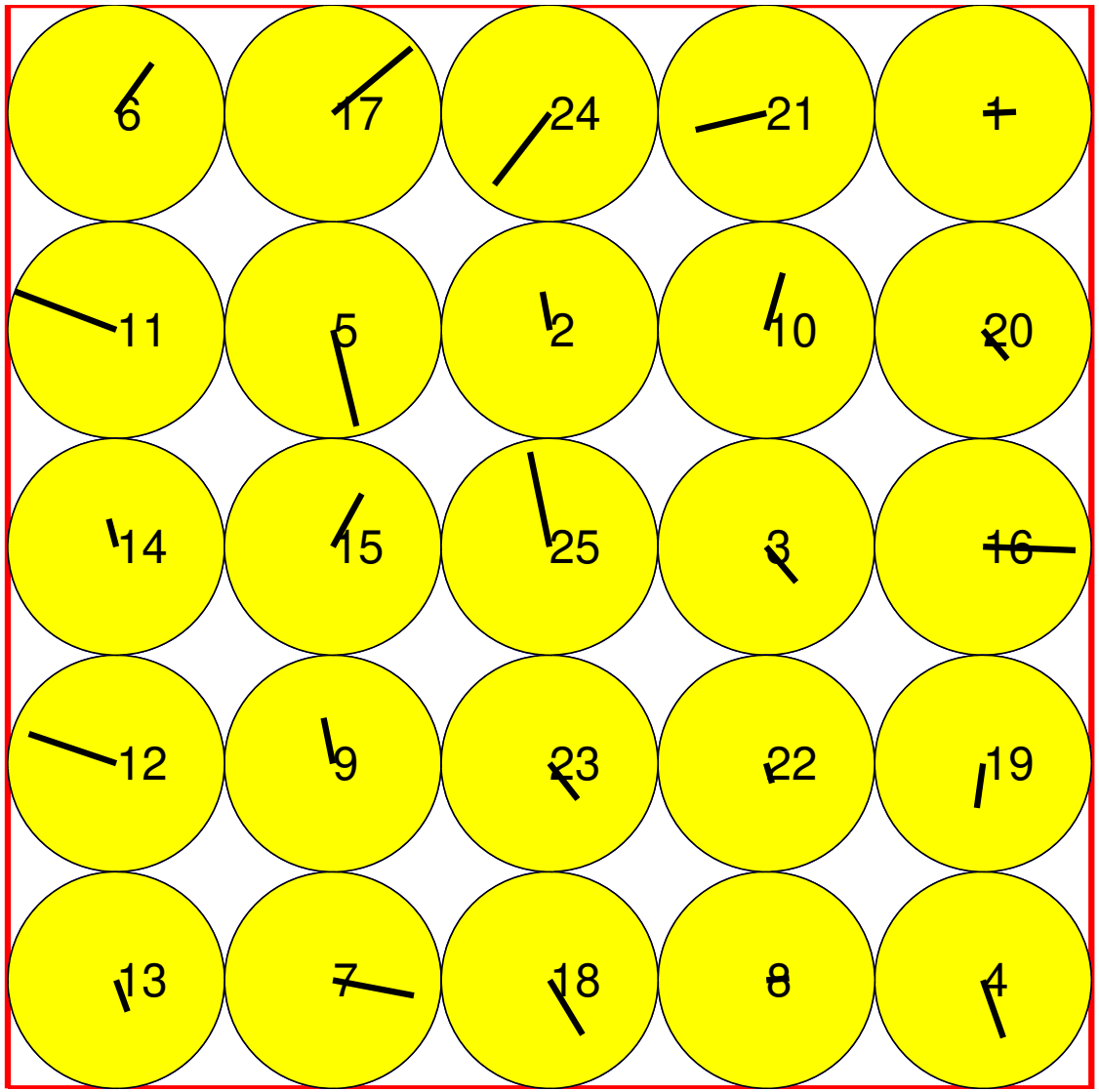
Rather efficient for $d = 2$ (if $R(t)$ increases slowly enough...)

but not very efficient for $d > 2$

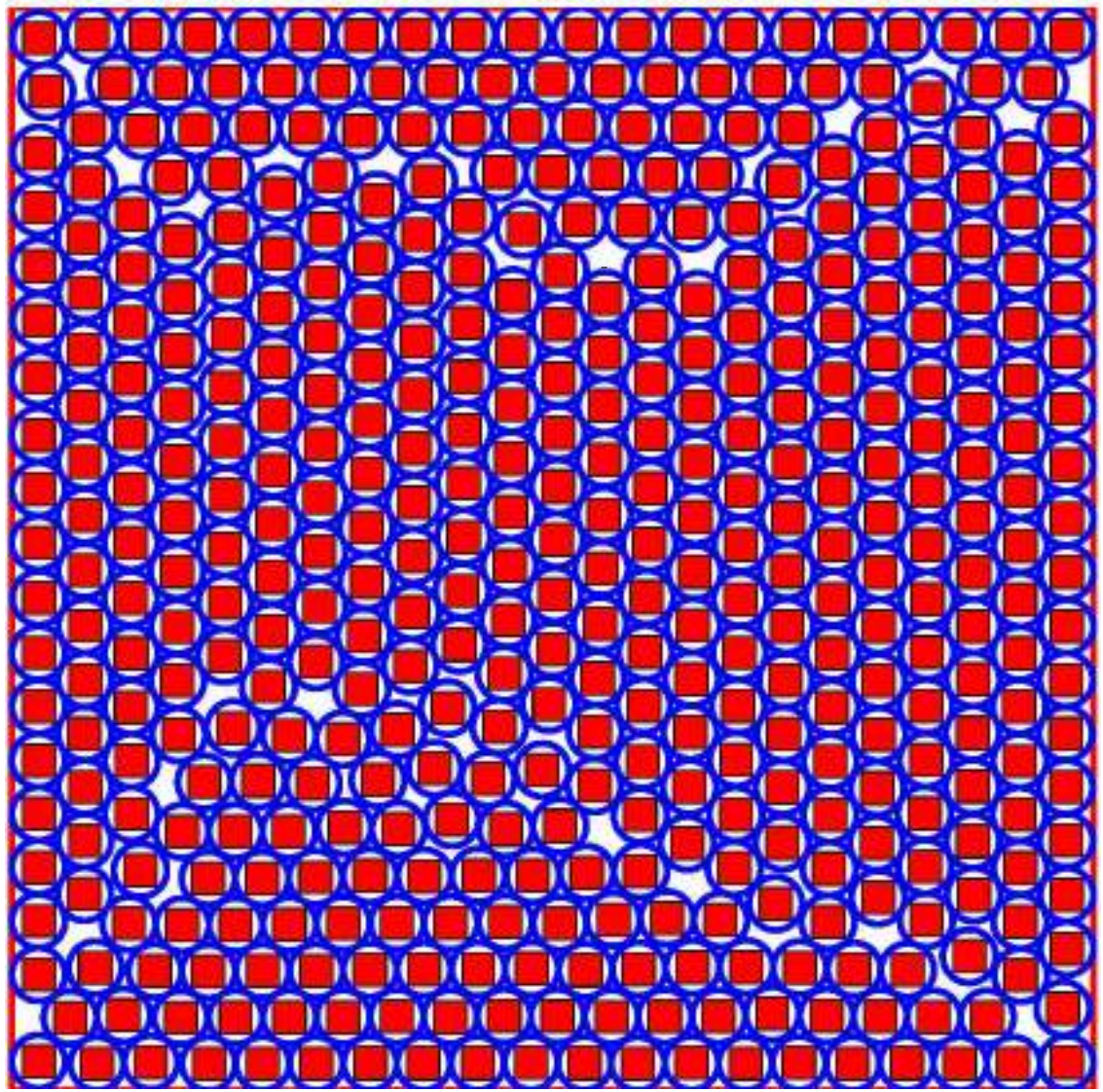
$n = 25$ ($R(t)$ increases too fast)



$n = 25$ ($R(t)$ increases slowly enough)



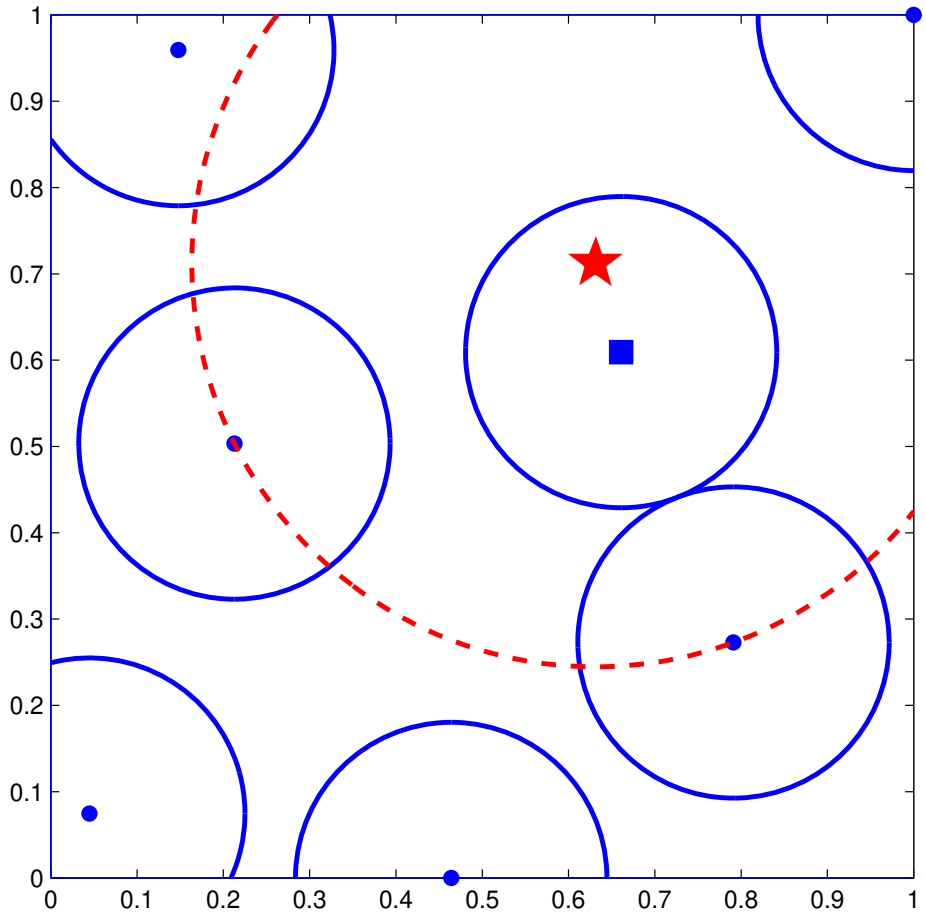
$n = 441$



2) **miniMax for Maximin**

Principle: repeat the following steps

a) Choose $\mathbf{x}_j$ from $\mathbf{X}_n$,
find $\mathbf{x}^*$ in $\mathscr{X}$ such that
$\min_{i \neq j} \|\mathbf{x}^* - \mathbf{x}_i\| =$
$\max_{\mathbf{x} \in \mathscr{X}} \min_{i \neq j} \|\mathbf{x} - \mathbf{x}_i\|$
(a byproduct of calculation of
$\Phi_{mM}(\mathbf{X}_{n \setminus j})$)



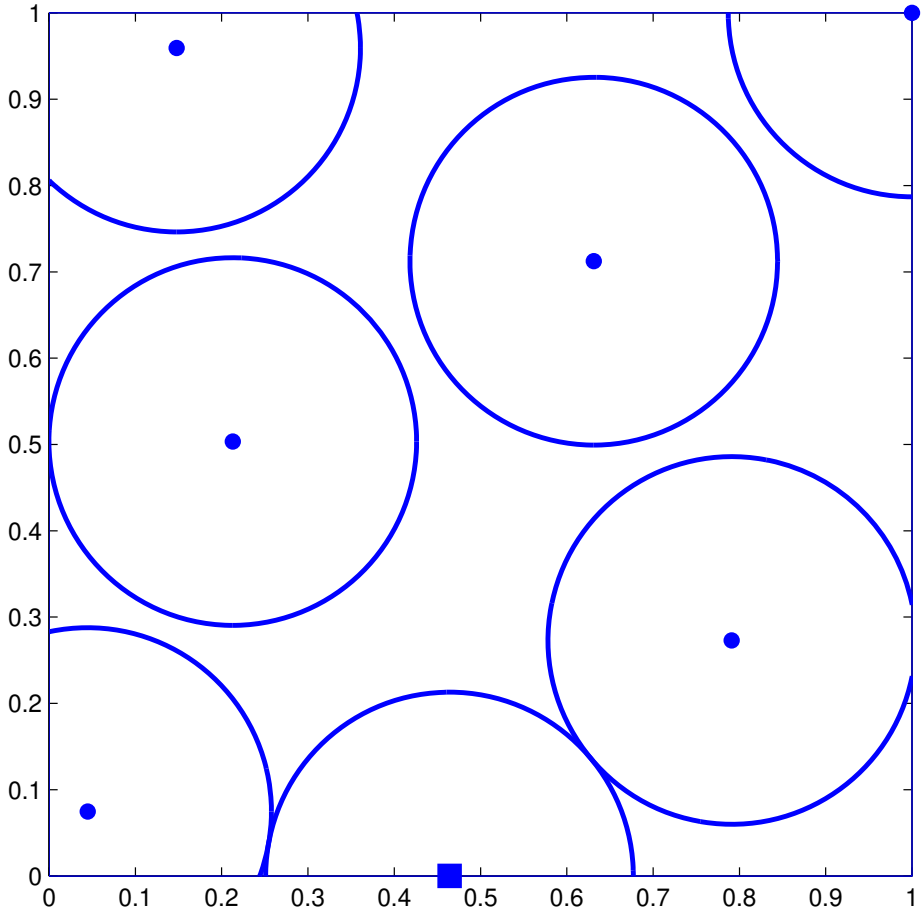
2) **miniMax for Maximin**

Principle: repeat the following steps

b) Replace $\mathbf{x}_j$ by $\mathbf{x}^*$

— which explains that

$$\Phi_{mM}(\mathbf{X}^*_{Mm,n}) \leq \Phi_{Mm}(\mathbf{X}^*_{Mm,n})$$



Related to "coffee-house design": start with $X_0 = \emptyset$, include points one by one $\mathbf{x}_1$ at the centre of $\mathscr{X}$, then $\mathbf{x}_{n+1}$ furthest point from X_n, $n \geq 1$
(called coffee-house design (Müller, 2007, Chap. 4))

Related to "coffee-house design": start with $X_0 = \emptyset$, include points one by one $\mathbf{x}_1$ at the centre of $\mathscr{X}$, then $\mathbf{x}_{n+1}$ furthest point from X_n, $n \geq 1$
(called coffee-house design (Müller, 2007, Chap. 4))

Guarantees $\boxed{\text{Eff}_{mM}(X_n) = \frac{\Phi^*_{mM,n}}{\Phi_{mM}(X_n)} \geq \frac{1}{2} \text{ and } \text{Eff}_{Mm}(X_n) = \frac{\Phi_{Mm}(X_n)}{\Phi^*_{Mm,n}} \geq \frac{1}{2} \text{ for all } n}$

with $\Phi_{Mm}(X_n) = \min_{i \neq j \in \{1,\ldots,n\}} \|\mathbf{x}_i - \mathbf{x}_j\|$ the maximin-distance criterion,
and $\Phi^*_{Mm,n}$ its optimal (maximum) value

Related to "coffee-house design": start with $X_0 = \emptyset$, include points one by one $\mathbf{x}_1$ at the centre of $\mathscr{X}$, then $\mathbf{x}_{n+1}$ furthest point from X_n, $n \geq 1$
(called coffee-house design (Müller, 2007, Chap. 4))

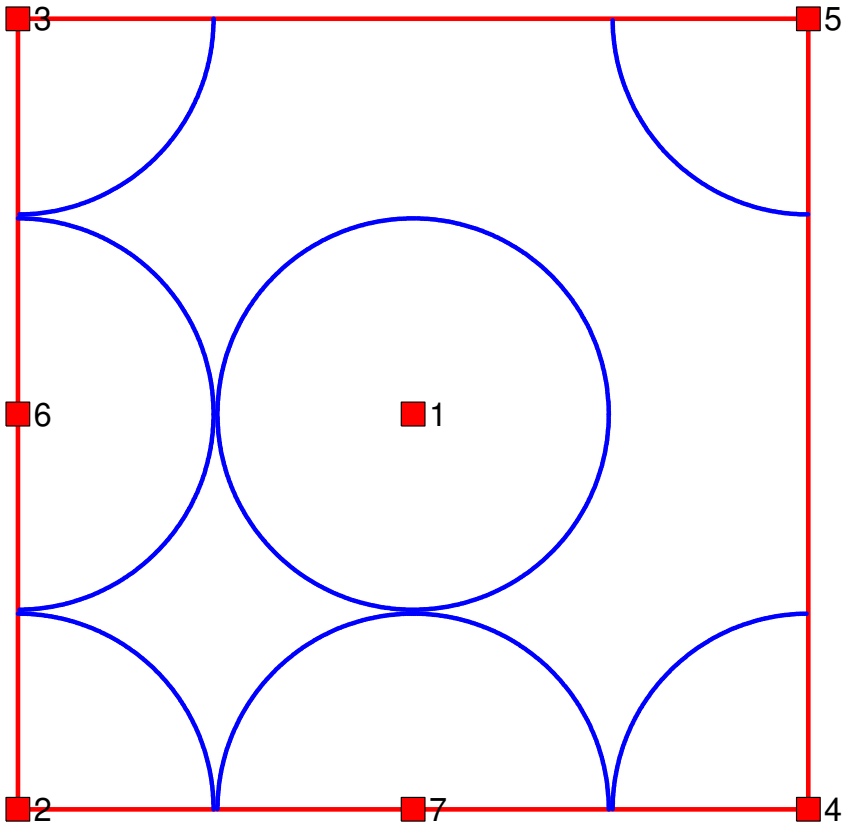
Guarantees $\boxed{\text{Eff}_{mM}(X_n) = \frac{\Phi^*_{mM,n}}{\Phi_{mM}(X_n)} \geq \frac{1}{2} \text{ and } \text{Eff}_{Mm}(X_n) = \frac{\Phi_{Mm}(X_n)}{\Phi^*_{Mm,n}} \geq \frac{1}{2} \text{ for all } n}$

with $\Phi_{Mm}(X_n) = \min_{i \neq j \in \{1,\ldots,n\}} \|\mathbf{x}_i - \mathbf{x}_j\|$ the maximin-distance criterion,
and $\Phi^*_{Mm,n}$ its optimal (maximum) value

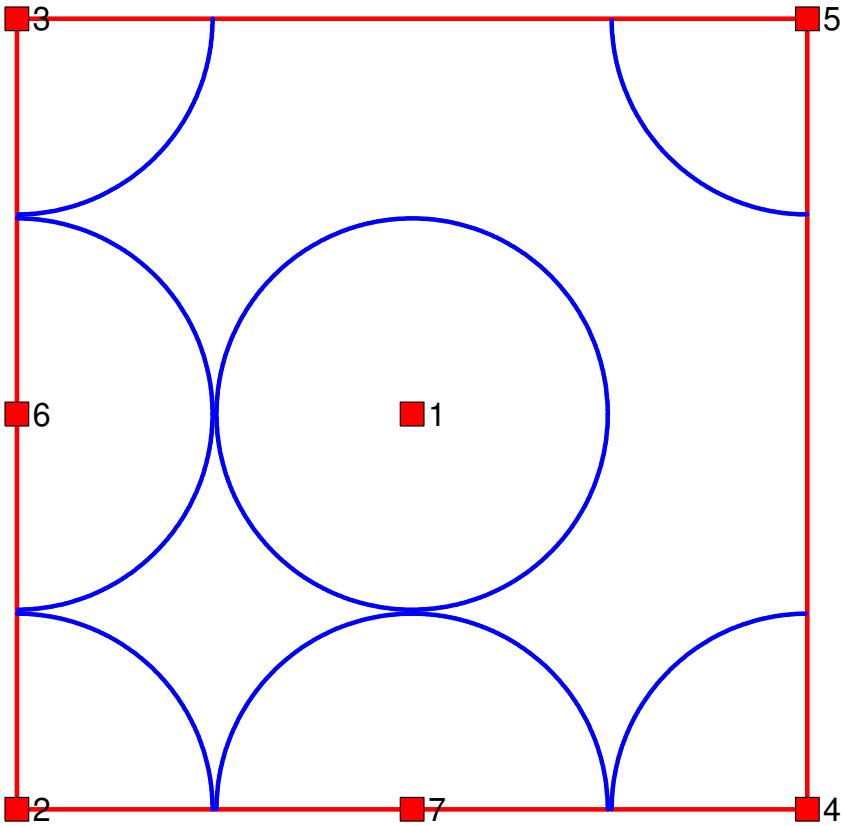
Proof. (Gonzalez, 1985) — repeated later

- by construction:
  $\Phi_{Mm}(X_{n+1}) \triangleq \min_{\mathbf{x}_i \neq \mathbf{x}_j \in X_{n+1}} \|\mathbf{x}_i - \mathbf{x}_j\| = d(\mathbf{x}_{n+1}, X_n) = \Phi_{mM}(X_n)$
- X^*_n a Φ_{mM}-optimal design: the n balls $\mathscr{B}(\mathbf{x}^*_i, \Phi_{mM}(X^*_n))$, $\mathbf{x}^*_i \in X^*_n$, cover $\mathscr{X}$
  $\Rightarrow$ one of them contains 2 points $\mathbf{x}_i$, $\mathbf{x}_j$ in X_{n+1} *for any* X_{n+1} *($n+1$ points)*
  $\Rightarrow \Phi_{Mm}(X_{n+1}) \leq \|\mathbf{x}_i - \mathbf{x}_j\| \leq 2\Phi_{mM}(X^*_n)$
  $\Rightarrow \Phi^*_{Mm,n+1} \leq 2\Phi_{mM}(X^*_n) \leq 2\Phi_{mM}(X_n) = \Phi_{Mm}(X_{n+1})$

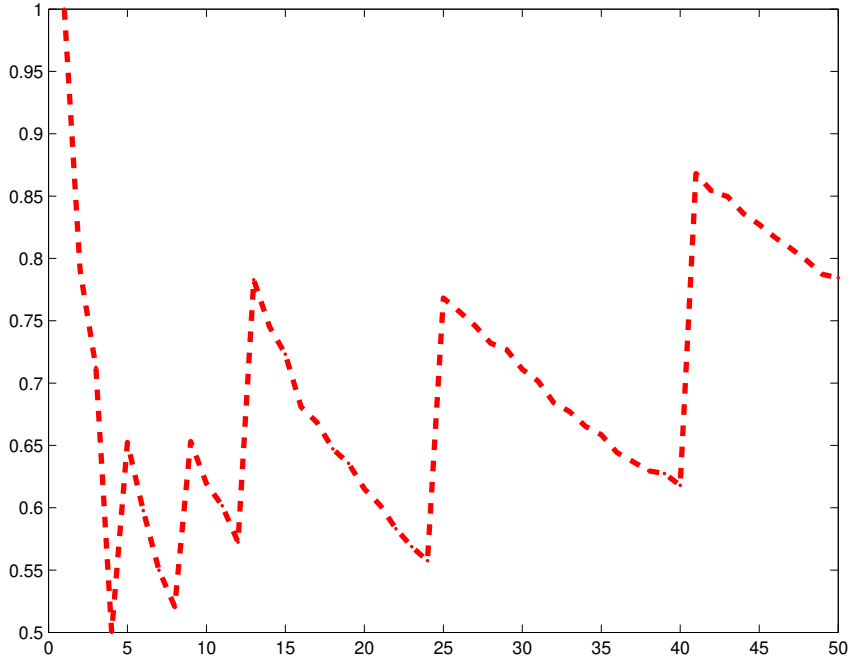
$\mathscr{X} = [0,1]^2,\ n = 7$



$\mathscr{X} = [0,1]^2$, $n = 7$



$\text{Eff}_{mM}(X_n)$, $n = 1 \ldots, 50$



Regular construction ➠ large fluctuations of $\text{Eff}_{mM}(X_n)$

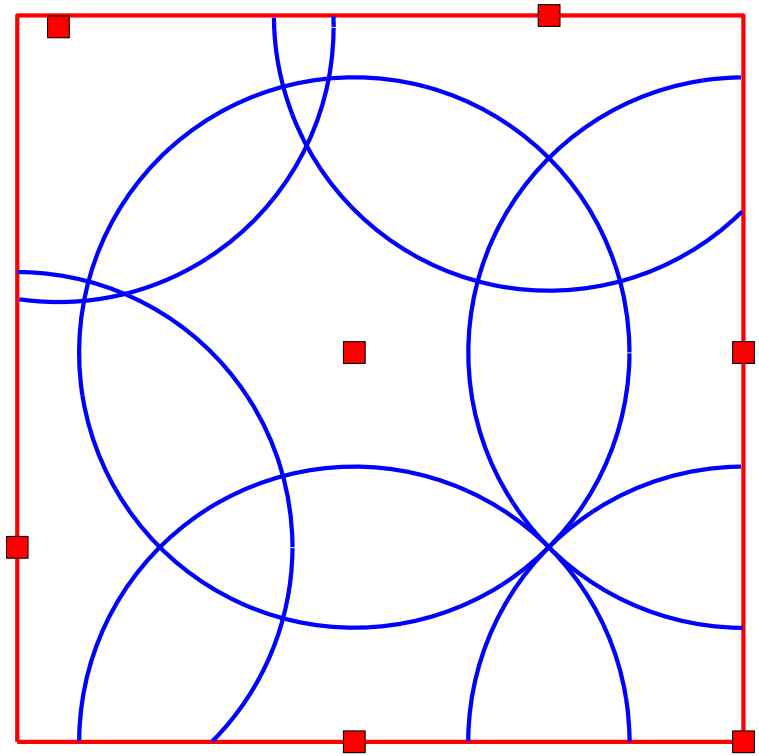
Remark: a Maximin distance design is useful for the miniMax criterion

Principle:

Points from a Maximin-optimal $\mathbf{X}^*_{Mm,n}$ tend to lie along the boundary of $\mathscr{X}$

➠ Apply an homothecy with center $\mathbf{c} \in \text{int}(\mathscr{X})$ and ratio $1/(1+\epsilon)$ to all points in $\mathbf{X}^*_{Mm,n}$ ($\mathscr{X} = [0,1]^d$, $\mathbf{c} = \frac{1}{2}\,\mathbf{1}$ ➠ $\mathbf{X}_n(\epsilon) = \mathbf{c} + \frac{1}{1+\epsilon}\,(\mathbf{X}^*_{Mm,n} - \mathbf{c})$)

$d = 2$, $n = 7$, $\mathbf{X}^*_{Mm,n}$ Maximin-optimal



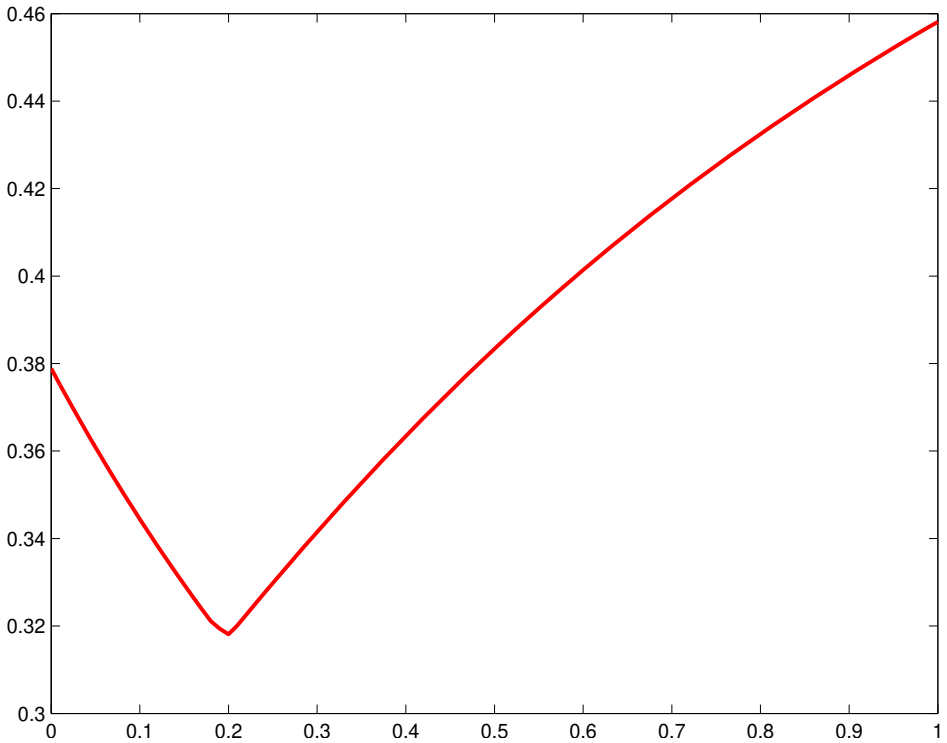
Remark: a Maximin distance design is useful for the miniMax criterion

Principle:

Points from a Maximin-optimal $\mathbf{X}^*_{Mm,n}$ tend to lie along the boundary of $\mathscr{X}$

➠ Apply an homothecy with center $\mathbf{c} \in \text{int}(\mathscr{X})$ and ratio $1/(1+\epsilon)$ to all points in $\mathbf{X}^*_{Mm,n}$ ($\mathscr{X} = [0,1]^d$, $\mathbf{c} = \frac{1}{2}\mathbf{1}$ ➠ $\mathbf{X}_n(\epsilon) = \mathbf{c} + \frac{1}{1+\epsilon}(\mathbf{X}^*_{Mm,n} - \mathbf{c})$)

$\Phi_{mM}(\mathbf{X}_n(\epsilon))$, $0 \leq \epsilon \leq 1$ ➠ $\epsilon^* = 0.2$



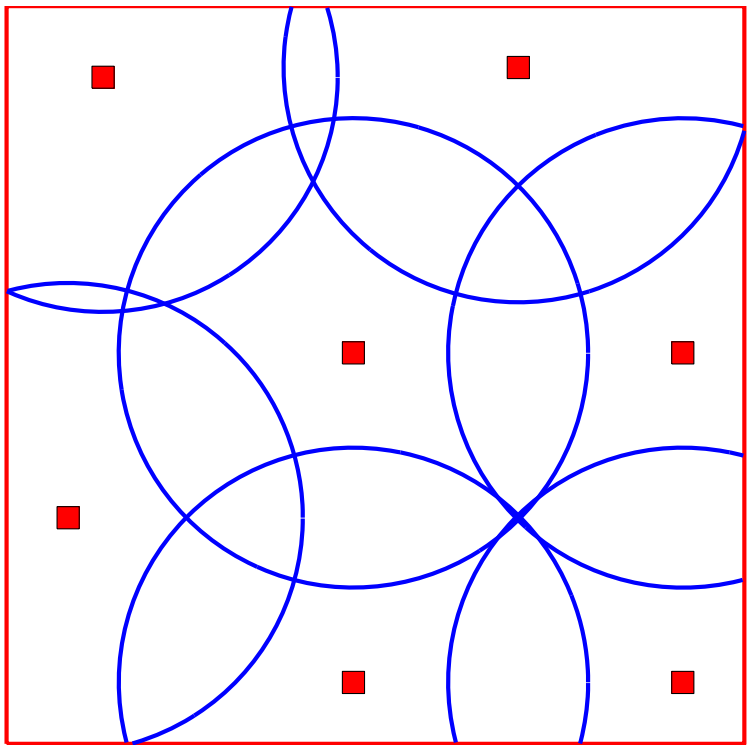
Remark: a Maximin distance design is useful for the miniMax criterion

Principle:

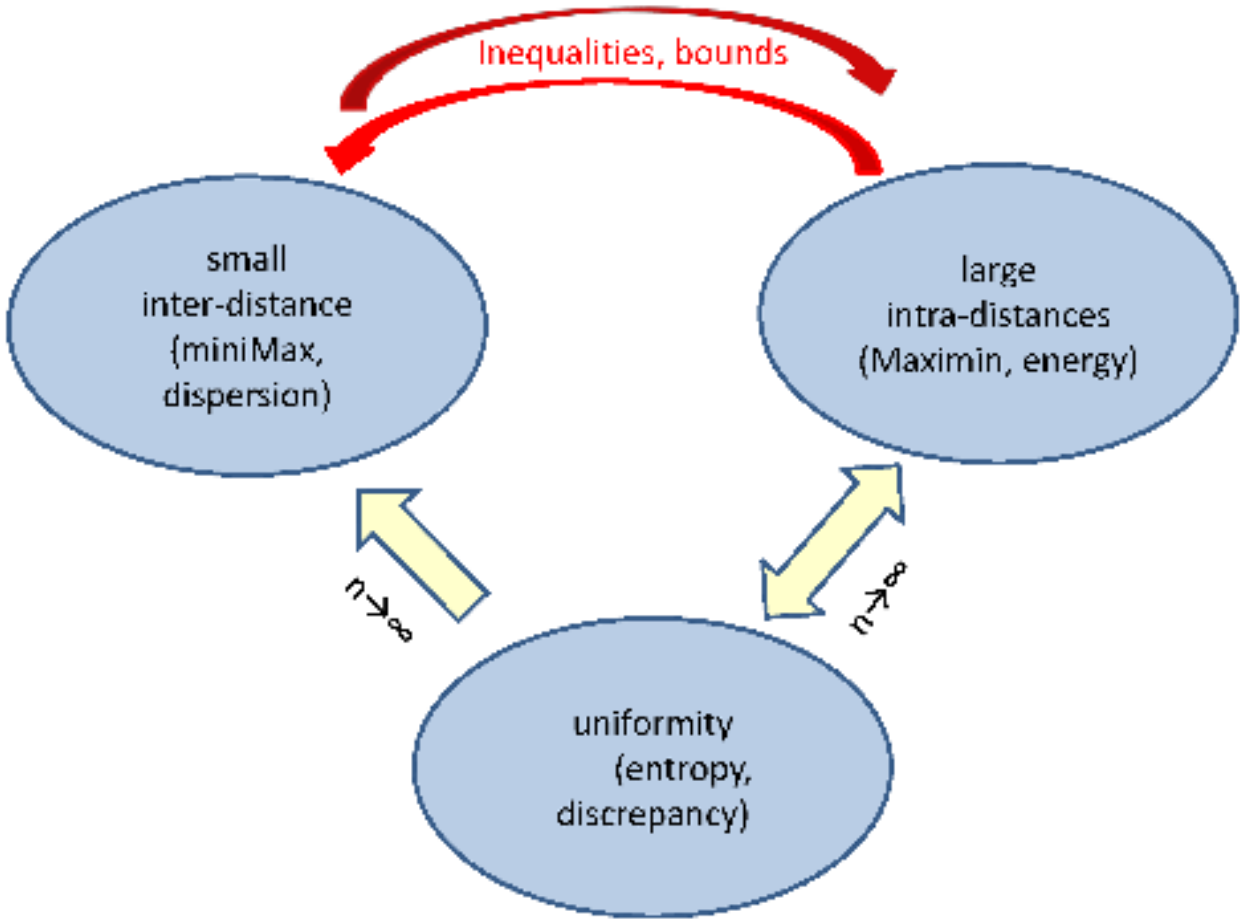
Points from a Maximin-optimal $\mathbf{X}^*_{Mm,n}$ tend to lie along the boundary of $\mathscr{X}$

➠ Apply an homothecy with center $\mathbf{c} \in \text{int}(\mathscr{X})$ and ratio $1/(1+\epsilon)$ to all points in $\mathbf{X}^*_{Mm,n}$ ($\mathscr{X} = [0,1]^d$, $\mathbf{c} = \frac{1}{2}\mathbf{1}$ ➠ $\mathbf{X}_n(\epsilon) = \mathbf{c} + \frac{1}{1+\epsilon}(\mathbf{X}^*_{Mm,n} - \mathbf{c})$)

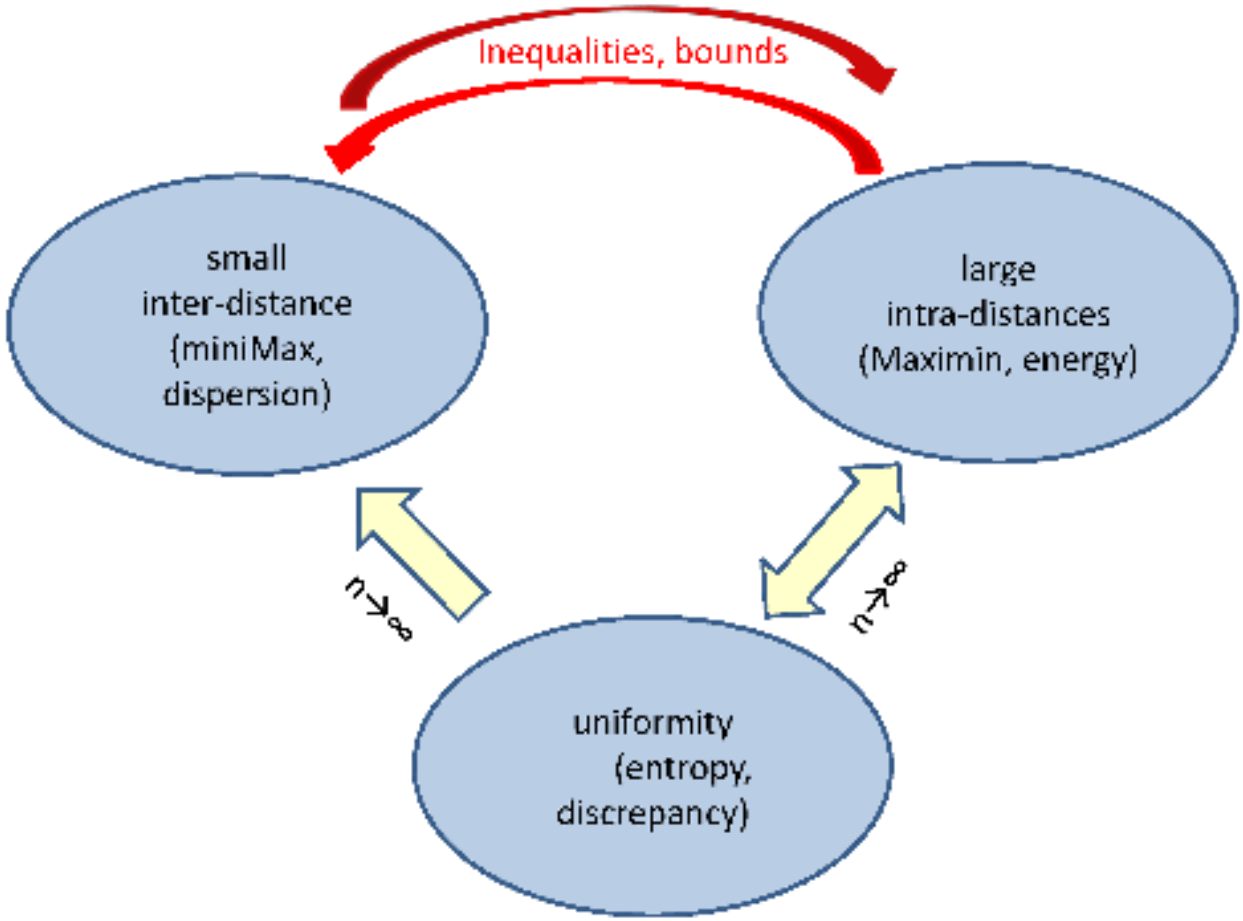
$\mathbf{X}_n(\epsilon^*)$ ➠ $\Phi_{mM}(\mathbf{X}_n(\epsilon^*)) = 0.3181$ (true miniMax optimum $= 0.2743$)



1.5 Relations between Φ_{Mm} and Φ_{mM} $(d \geq 2)$



1.5 Relations between Φ_{Mm} and Φ_{mM} $(d \geq 2)$



Regularized Maximin

Notation: $\mathscr{X} = [0,1]^d$, $V_d = \text{vol}[\mathscr{B}(\mathbf{0},1)] = \pi^{d/2}/\Gamma(d/2+1)$

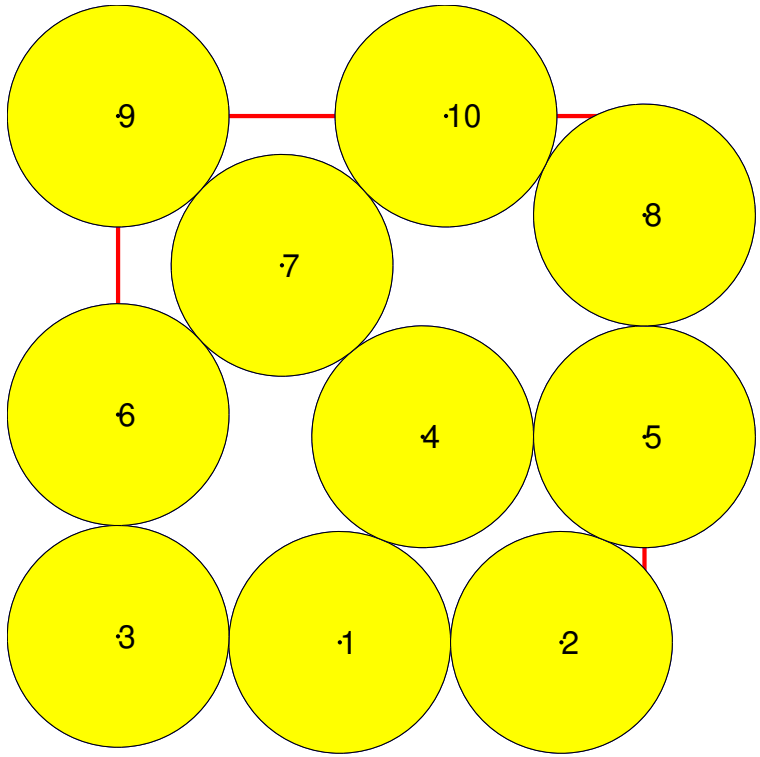
$\triangleright$ $\Phi^*_{mM,n} \triangleq \min_{\mathbf{X}_n} \Phi_{mM}(\mathbf{X}_n)$, $\Phi^*_{Mm,n} \triangleq \max_{\mathbf{X}_n} \Phi_{Mm}(\mathbf{X}_n)$ $\triangleleft$

Notation: $\mathscr{X} = [0,1]^d$, $V_d = \text{vol}[\mathscr{B}(\mathbf{0},1)] = \pi^{d/2}/\Gamma(d/2+1)$

$\triangleright$ $\Phi^*_{mM,n} \triangleq \min_{\mathbf{X}_n} \Phi_{mM}(\mathbf{X}_n)$, $\Phi^*_{Mm,n} \triangleq \max_{\mathbf{X}_n} \Phi_{Mm}(\mathbf{X}_n)$ $\triangleleft$

- $\boxed{\frac{1}{2}\Phi_{Mm}(\mathbf{X}_n) \leq \Phi_{mM}(\mathbf{X}_n), \text{ for all } \mathbf{X}_n \ (n \geq 2)}$

(the n balls $\mathscr{B}(\mathbf{x}_i, \frac{1}{2}\Phi_{Mm}(\mathbf{X}_n))$ do not cover $\mathscr{X}$)



- $\boxed{\frac{1}{2}\Phi_{Mm}(\mathbf{X}_{n+1}) \leq \Phi^*_{mM,n}, \text{ for all } \mathbf{X}_{n+1} \ (n \geq 1)}$

Proof: one of the n balls $\mathscr{B}(\mathbf{z}_i, \Phi^*_{mM,n})$, $\mathbf{z}_i \in \mathbf{X}^*_{mM,n}$, contains 2 points $\mathbf{x}_i$ and $\mathbf{x}_j$ from $\mathbf{X}_{n+1}$

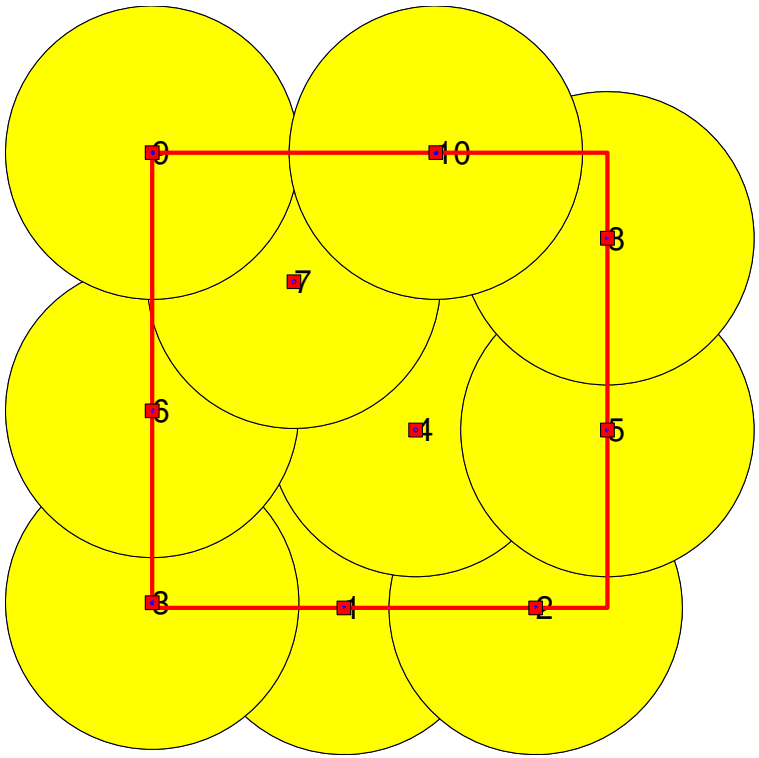
$$\Longrightarrow \quad \Phi_{Mm}(\mathbf{X}_{n+1}) \leq \|\mathbf{x}_i - \mathbf{x}_j\| \leq 2\Phi^*_{mM,n}$$

- $\boxed{\frac{1}{2}\Phi_{Mm}(\mathbf{X}_{n+1}) \leq \Phi^*_{mM,n}, \text{ for all } \mathbf{X}_{n+1} \ (n \geq 1)}$

  Proof: one of the n balls $\mathscr{B}(\mathbf{z}_i, \Phi^*_{mM,n})$, $\mathbf{z}_i \in \mathbf{X}^*_{mM,n}$, contains 2 points $\mathbf{x}_i$ and $\mathbf{x}_j$ from $\mathbf{X}_{n+1}$

$$\Longrightarrow \quad \Phi_{Mm}(\mathbf{X}_{n+1}) \leq \|\mathbf{x}_i - \mathbf{x}_j\| \leq 2\Phi^*_{mM,n}$$

- The n balls $\mathscr{B}(\mathbf{x}_i, \Phi_{mM}(\mathbf{X}_n))$ cover $\mathscr{X}$, for all $\mathbf{X}_n$



- Sphere covering $\Rightarrow nV_d \left(\Phi^*_{mM,n}\right)^d > 1$

$$\boxed{R^*_n < \Phi^*_{mM,n}} \quad \text{with } R^*_n = (nV_d)^{-1/d}$$

- Sphere covering $\Rightarrow nV_d \left(\Phi^*_{mM,n}\right)^d > 1$
- $\Phi^*_{mM,n} \leq \boxed{\Phi_{mM}(\mathbf{X}^*_{Mm,n}) \leq \Phi_{Mm}(\mathbf{X}^*_{Mm,n})} = \Phi^*_{Mm,n}$

  (proof by contradiction)

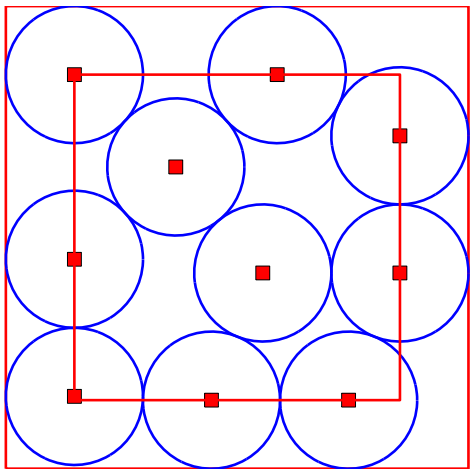
$$\boxed{R^*_n < \Phi^*_{mM,n} \leq \Phi^*_{Mm,n}} \quad \text{with } R^*_n = (nV_d)^{-1/d}$$

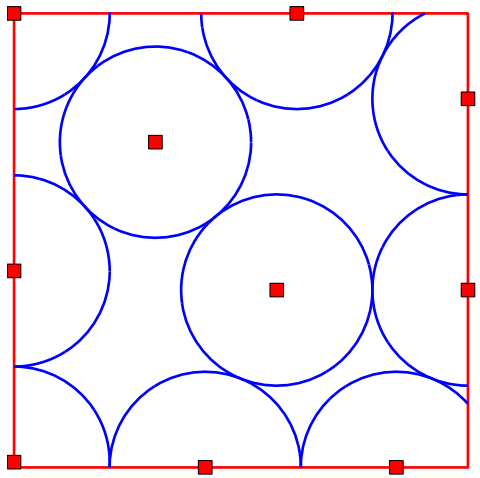
- Sphere covering $\Rightarrow nV_d\,(\Phi^*_{mM,n})^d > 1$
- $\Phi^*_{mM,n} \leq \boxed{\Phi_{mM}(\mathbf{X}^*_{Mm,n}) \leq \Phi_{Mm}(\mathbf{X}^*_{Mm,n})} = \Phi^*_{Mm,n}$

  (proof by contradiction)
- packing of n balls with radius R in $[0,1]^d$

  $\Rightarrow nV_d\,R^d < 1$, i.e., $\boxed{R < R^*_n}$

  (and $R^*_n < \frac{1}{2}$ for $n > \lceil 2^d/V_d \rceil$)





$$\boxed{R^*_n < \Phi^*_{mM,n} \leq \Phi^*_{Mm,n} < \frac{2R^*_n}{1-2R^*_n}} \quad \text{with } R^*_n = (nV_d)^{-1/d}$$

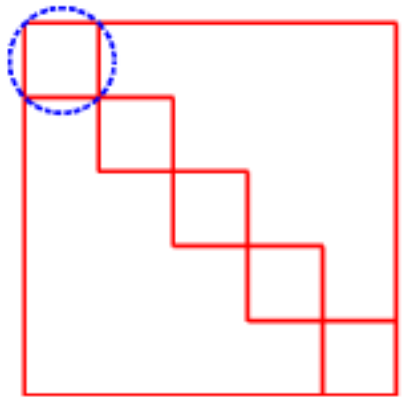
(these bounds are rather loose — factor $\frac{2}{1-2R^*_n} > 2$)

Other upper bounds on $\Phi^*_{mM,n} = \Phi_{mM}(X^*_n)$ when $\mathscr{X} = [0,1]^d$

Upper bound: use any design!

m^d-point regular grid in $\mathscr{X}$:

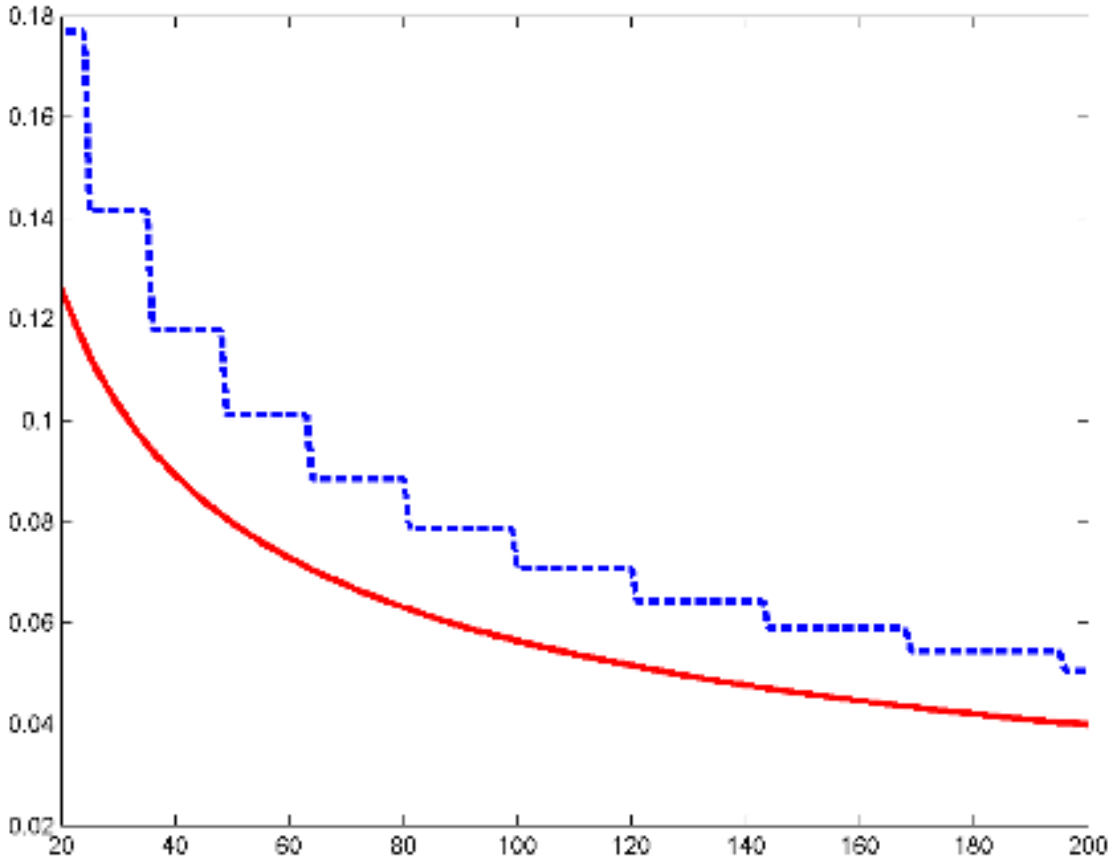
$$\Phi^*_{mM,m^d} \leq \frac{\sqrt{d}}{2m}:$$



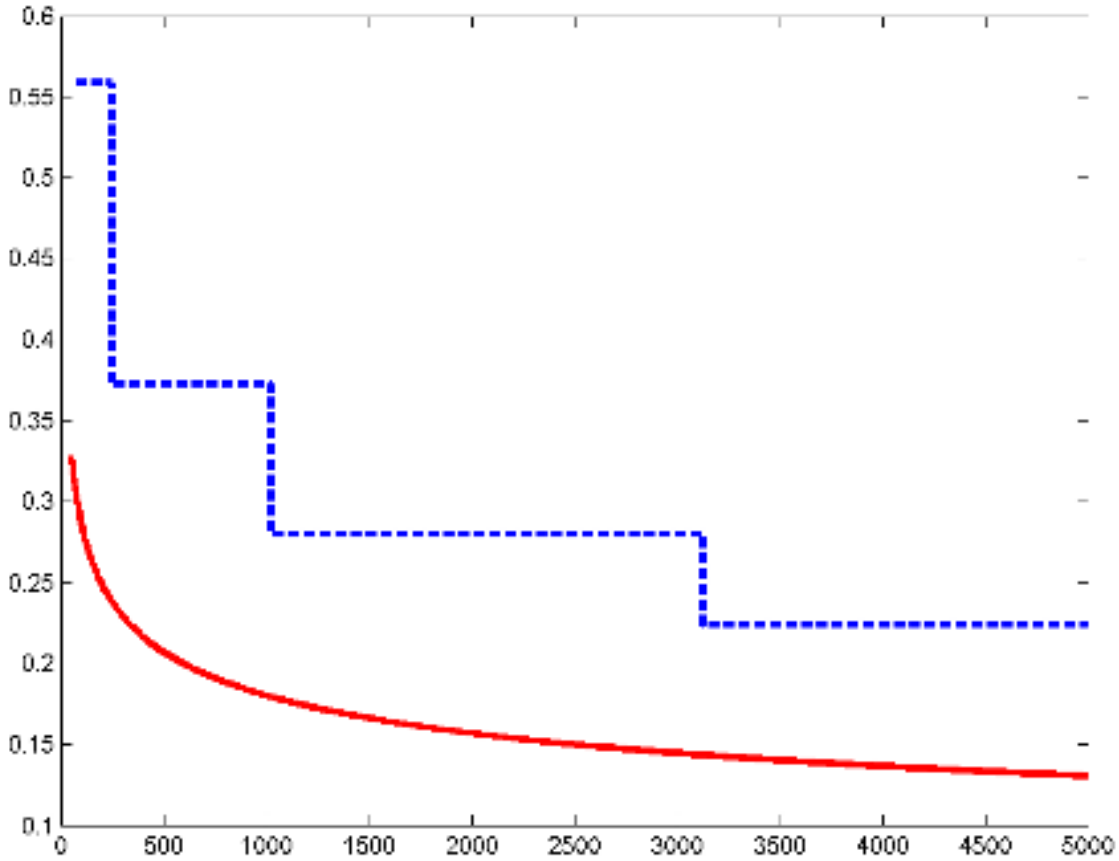
Take $m = \lfloor n^{1/d} \rfloor$, so that $m^d \leq n$ and $\Phi^*_{mM,n} \leq \Phi^*_{mM,m^d}$, therefore

$$\boxed{\Phi^*_{mM,n} \leq \overline{R}^*_n = \frac{\sqrt{d}}{2\lfloor n^{1/d} \rfloor}}$$

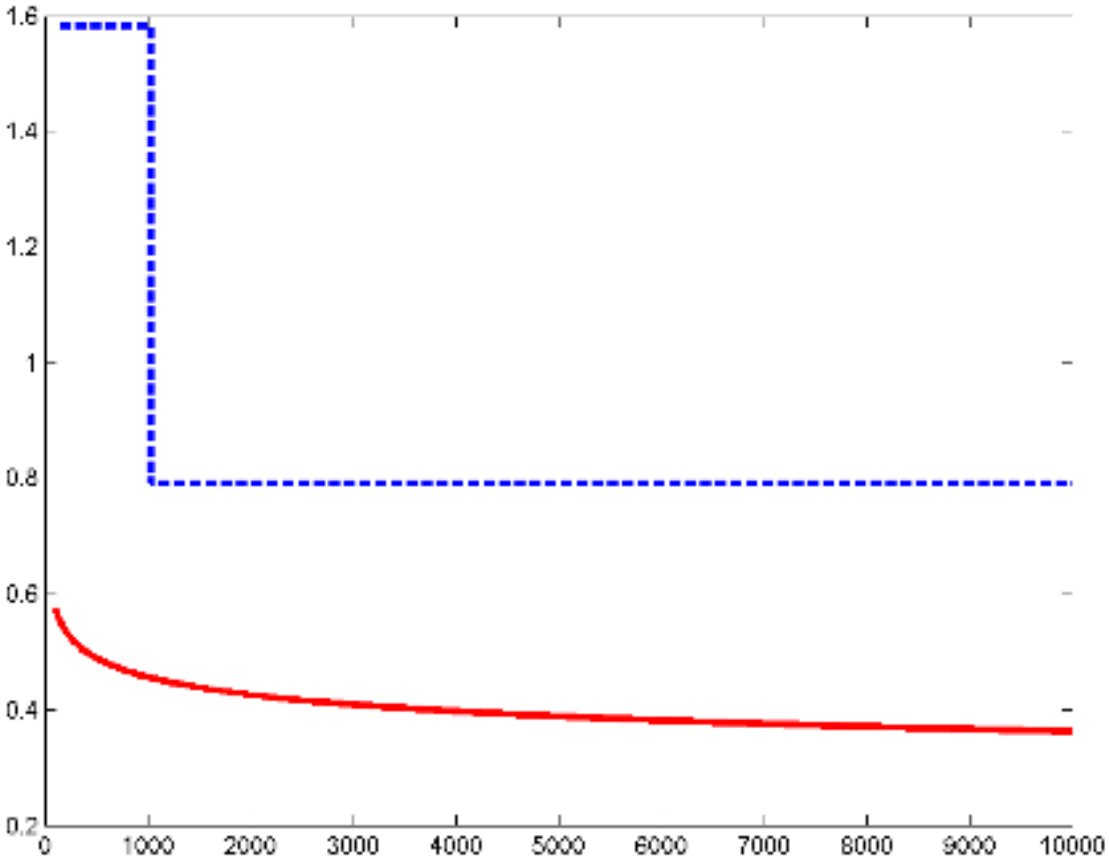
$$d = 2,\ R_n^* \leq \Phi_{mM,n}^* \leq \overline{R}_n^*$$



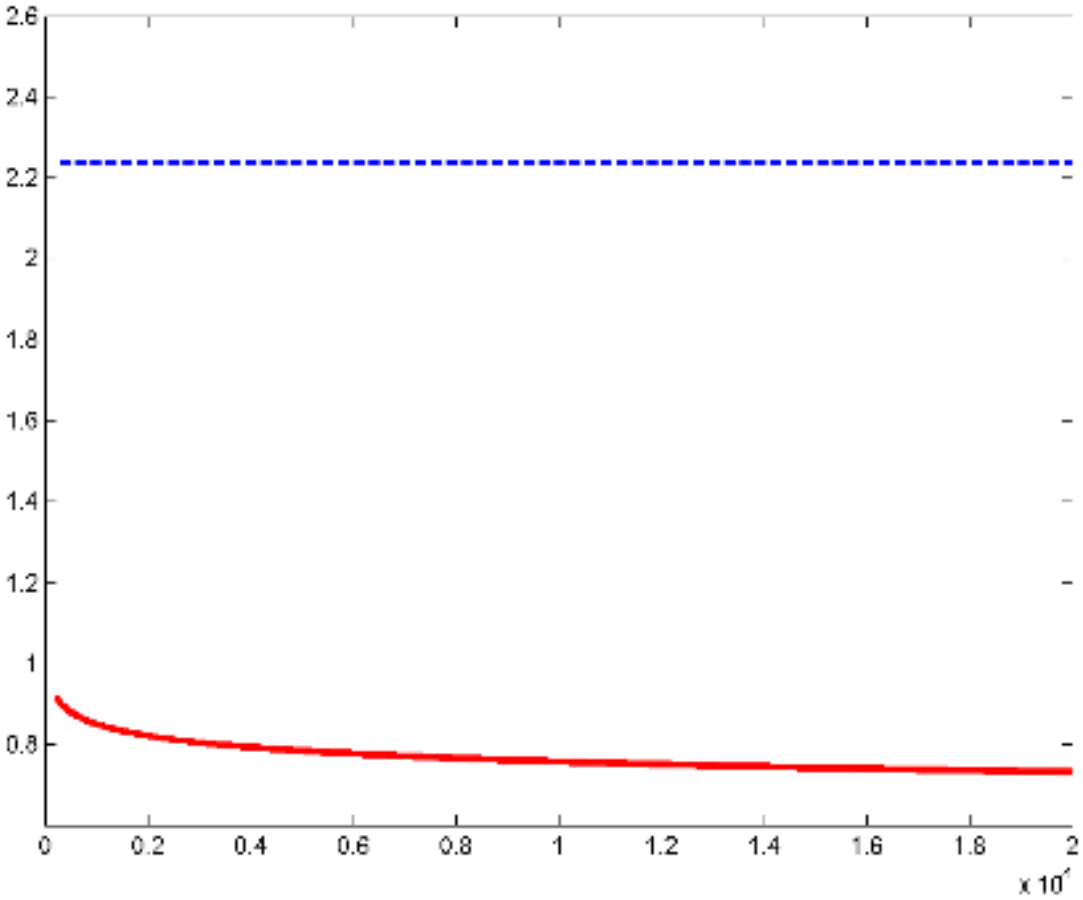
$$d = 5,\ R_n^* \leq \Phi_{mM,n}^* \leq \overline{R}_n^*$$



$$d = 10,\ R_n^* \leq \Phi_{mM,n}^* \leq \overline{R}_n^*$$



$$d = 20,\ R_n^* \leq \Phi_{mM,n}^* \leq \overline{R}_n^*$$



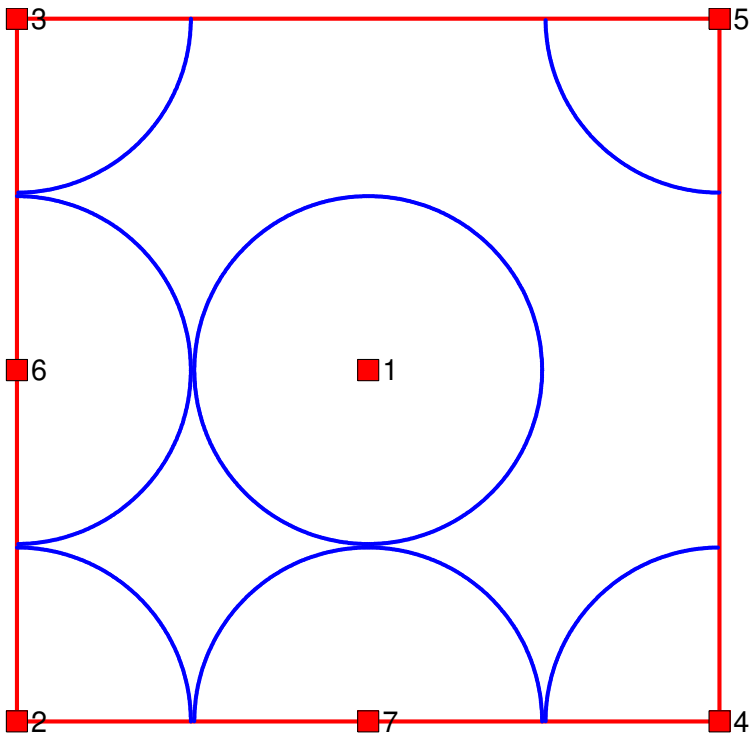
Why are such bounds useful?

- They give an idea of the suboptimality of a given design (for small d)

Why are such bounds useful?

- They give an idea of the suboptimality of a given design (for small d)
- They help understand algorithms:
  - 1/ Method "miniMax for Maximin"

    uses $\boxed{\Phi_{mM}(\mathbf{X}^*_{Mm,n}) \leq \Phi_{Mm}(\mathbf{X}^*_{Mm,n})}$
  - 2/ **Greedy algorithm of "coffee-house design"**
    0) Choose $\mathbf{x}_1 \in \mathscr{X}$, set $\mathbf{X}_1 = \{\mathbf{x}_1\}$
    1) For $k = 1, 2, \ldots$, find $\mathbf{x}^* = \arg\max_{\mathbf{x} \in \mathscr{X}} d(\mathbf{x}, \mathbf{X}_k)$, set $\mathbf{X}_{k+1} = \mathbf{X}_k \cup \{\mathbf{x}^*\}$



A bit primitive . . . but (Gonzalez, 1985) :

$$\frac{\Phi^*_{mM,k}}{\Phi_{mM}(X_k)} \geq \frac{1}{2} \ (k \geq 1) \quad \text{and} \quad \frac{\Phi_{Mm}(X_k)}{\Phi^*_{Mm,k}} \geq \frac{1}{2} \ (k \geq 2)$$

A bit primitive . . . but (Gonzalez, 1985) :

$$\frac{\Phi^*_{mM,k}}{\Phi_{mM}(X_k)} \geq \frac{1}{2} \ (k \geq 1) \quad \text{and} \quad \frac{\Phi_{Mm}(X_k)}{\Phi^*_{Mm,k}} \geq \frac{1}{2} \ (k \geq 2)$$

Proof: par construction, $\Phi_{Mm}(X_{k+1}) = \Phi_{mM}(X_k)$ for all $k \geq 1$

$\boxed{\frac{1}{2}\Phi_{Mm}(\mathbf{X}_{k+1}) \leq \Phi^*_{mM,k}, \text{ for all } \mathbf{X}_{k+1} \ (k \geq 1)}$ implies

a) $\Phi^*_{mM,k} \geq (1/2)\,\Phi_{mM}(X_k)$ and
b) $\Phi^*_{Mm,k+1} \leq 2\Phi^*_{mM,k} \leq 2\,\Phi_{mM}(X_k) = 2\,\Phi_{Mm}(X_{k+1})$

⏩ Regularized Maximin

Remark:

- There exist better bounds $d = 2, 3$: packing n balls with radius R in $[0, 1]^d$ $\Rightarrow nV_d R^d < \delta_d =$ *packing density*, with
  - $\delta_2 = \pi\sqrt{3}/6 \simeq 0.9069$
    [Lagrange, 1773 for lattices, Tóth 1940 for general packings]
  - $\delta_3 = \pi\sqrt{2}/6 \simeq 0.7405$
    [Kepler conjecture 1611, Gauss 1831 for lattices, Hales-Ferguson 2006 for general packings]
- Little is known for $d > 3$: best (densest) lattice packing known for $d \leq 8$, best general packings are known for $d = 8$ (Viazovska, 2016) and $d = 24$ (Cohn et al., 2017)

Remark:

- There exist better bounds $d = 2, 3$: packing n balls with radius R in $[0,1]^d$ $\Rightarrow nV_d\, R^d < \delta_d = \textit{packing density}$, with
  - $\delta_2 = \pi\sqrt{3}/6 \simeq 0.9069$
    [Lagrange, 1773 for lattices, Tóth 1940 for general packings]
  - $\delta_3 = \pi\sqrt{2}/6 \simeq 0.7405$
    [Kepler conjecture 1611, Gauss 1831 for lattices, Hales-Ferguson 2006 for general packings]
- Little is known for $d > 3$: best (densest) lattice packing known for $d \leq 8$, best general packings are known for $d = 8$ (Viazovska, 2016) and $d = 24$ (Cohn et al., 2017)
- There exist recent results (Wahl et al., 2014) on the distribution of $\Phi_{Mm}(\mathbf{X}_n)$ when the $\mathbf{x}_i$ are i.i.d. uniformly in $[0,1]^d$
  (but all $\|\mathbf{x}_i - \mathbf{x}_j\|^2$ tend to concentrate around $d/6$ when d gets large)

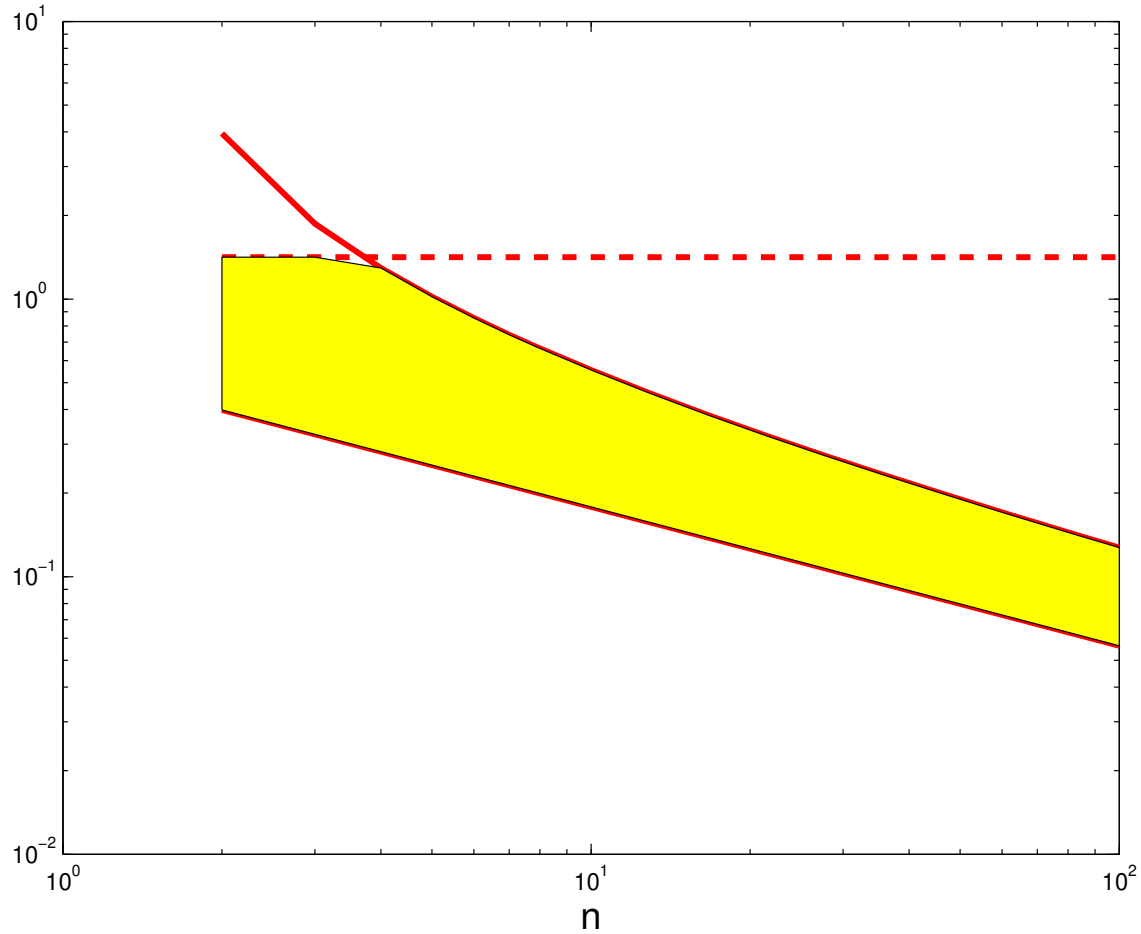
Remark:

- There exist better bounds $d = 2, 3$: packing n balls with radius R in $[0,1]^d$ $\Rightarrow nV_d\, R^d < \delta_d = \textit{packing density}$, with
  - $\delta_2 = \pi\sqrt{3}/6 \simeq 0.9069$
    [Lagrange, 1773 for lattices, Tóth 1940 for general packings]
  - $\delta_3 = \pi\sqrt{2}/6 \simeq 0.7405$
    [Kepler conjecture 1611, Gauss 1831 for lattices, Hales-Ferguson 2006 for general packings]
- Little is known for $d > 3$: best (densest) lattice packing known for $d \leq 8$, best general packings are known for $d = 8$ (Viazovska, 2016) and $d = 24$ (Cohn et al., 2017)
- There exist recent results (Wahl et al., 2014) on the distribution of $\Phi_{Mm}(\mathbf{X}_n)$ when the $\mathbf{x}_i$ are i.i.d. uniformly in $[0,1]^d$
  (but all $\|\mathbf{x}_i - \mathbf{x}_j\|^2$ tend to concentrate around $d/6$ when d gets large)

⏩ Regularized Maximin

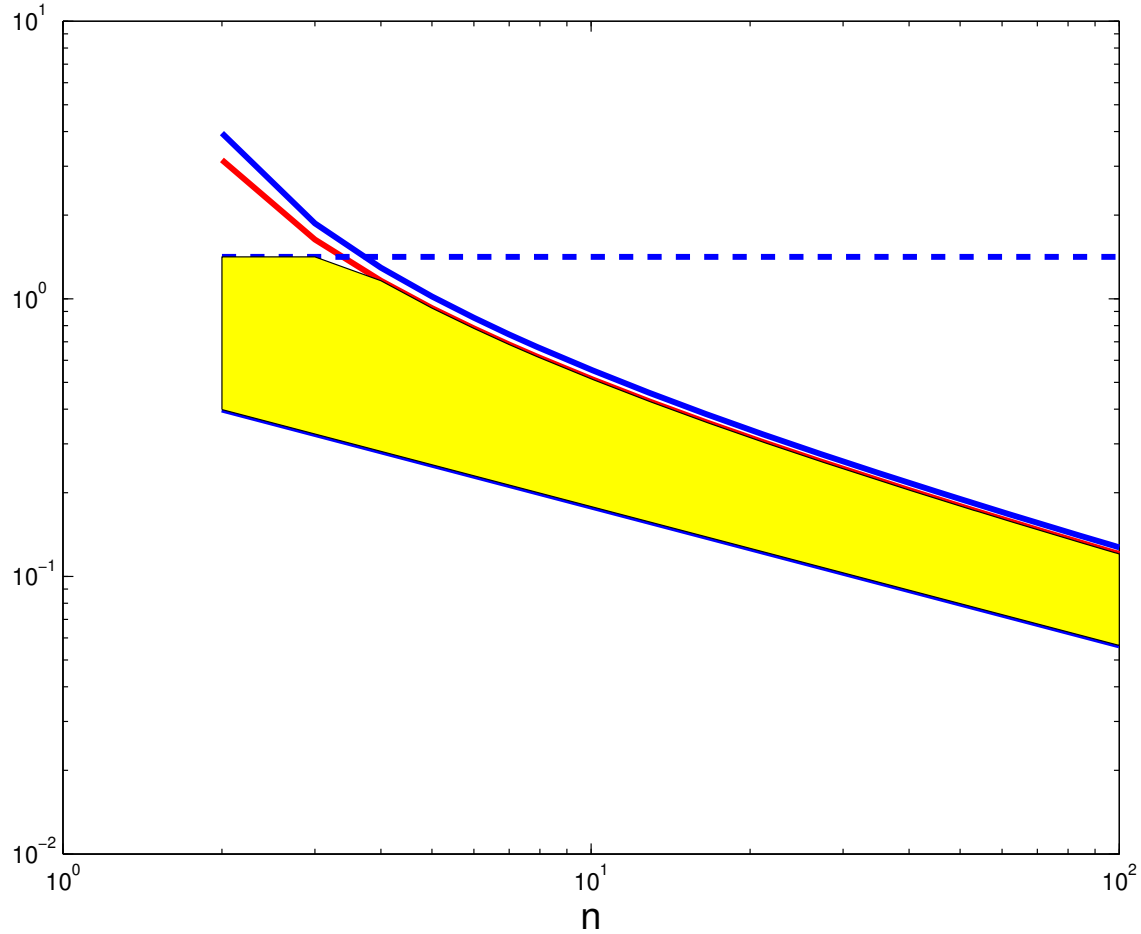
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 2$

$d = 2$ using $\boxed{R^*_n < \Phi^*_{mM,n} \leq \Phi^*_{Mm,n} < \min\{\frac{2R^*_n}{1-2R^*_n}, \sqrt{d}\}}$



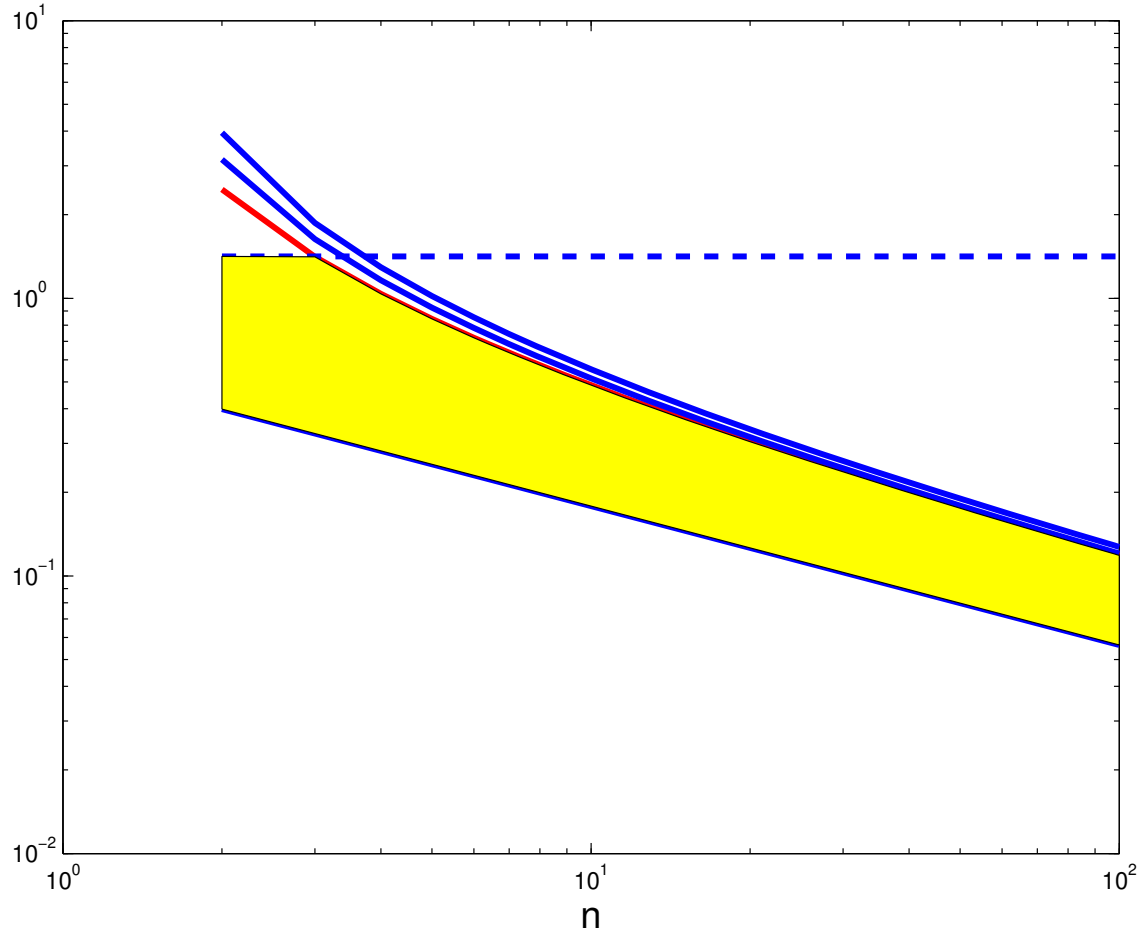
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 2$

$d = 2$ with moreover $\boxed{\Phi^*_{Mm,n} < \frac{1}{3^{1/4}\sqrt{n/2-1}}}$ (*packing density*)



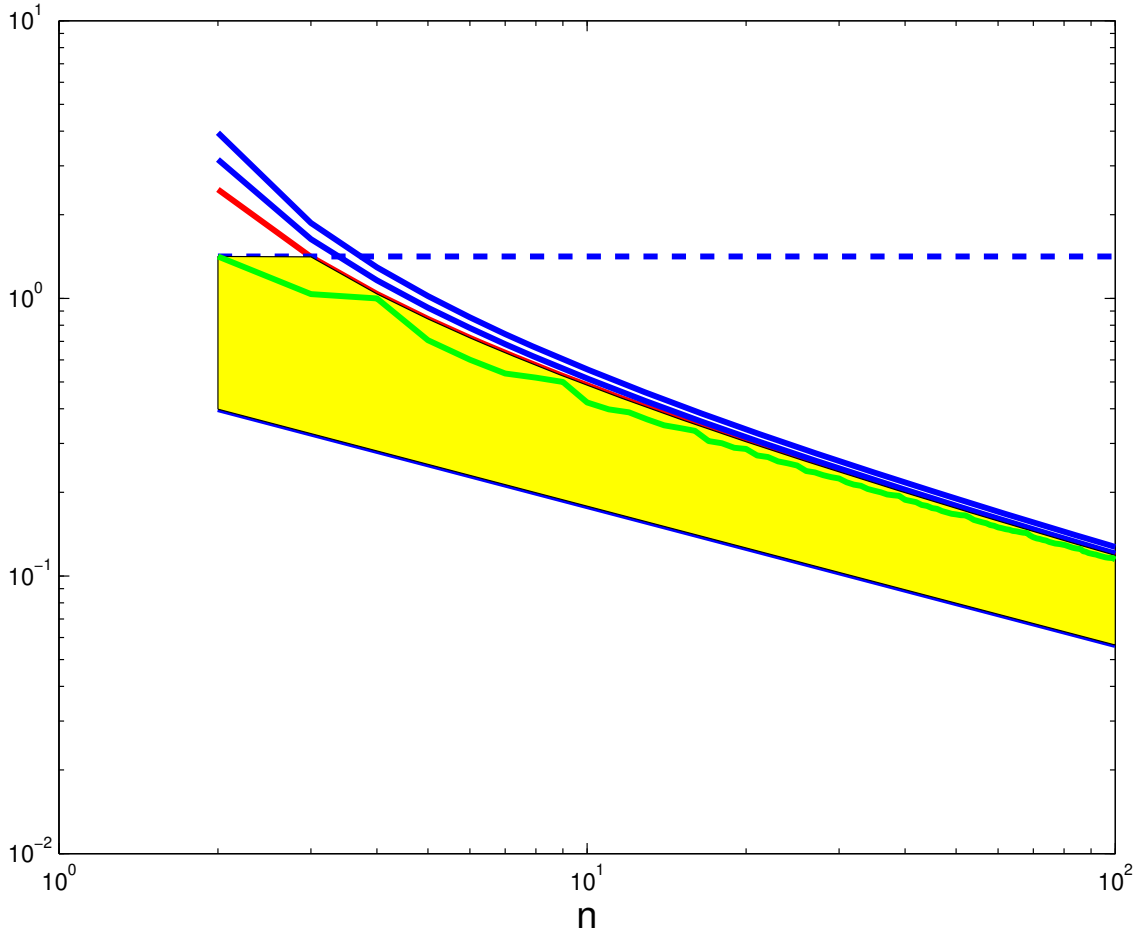
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 2$

$d = 2$ with moreover $\boxed{\Phi^*_{Mm,n} < \frac{1+\sqrt{1+2\,(n-1)\,/\sqrt{3}}}{n-1}}$ (Oler, 1961)



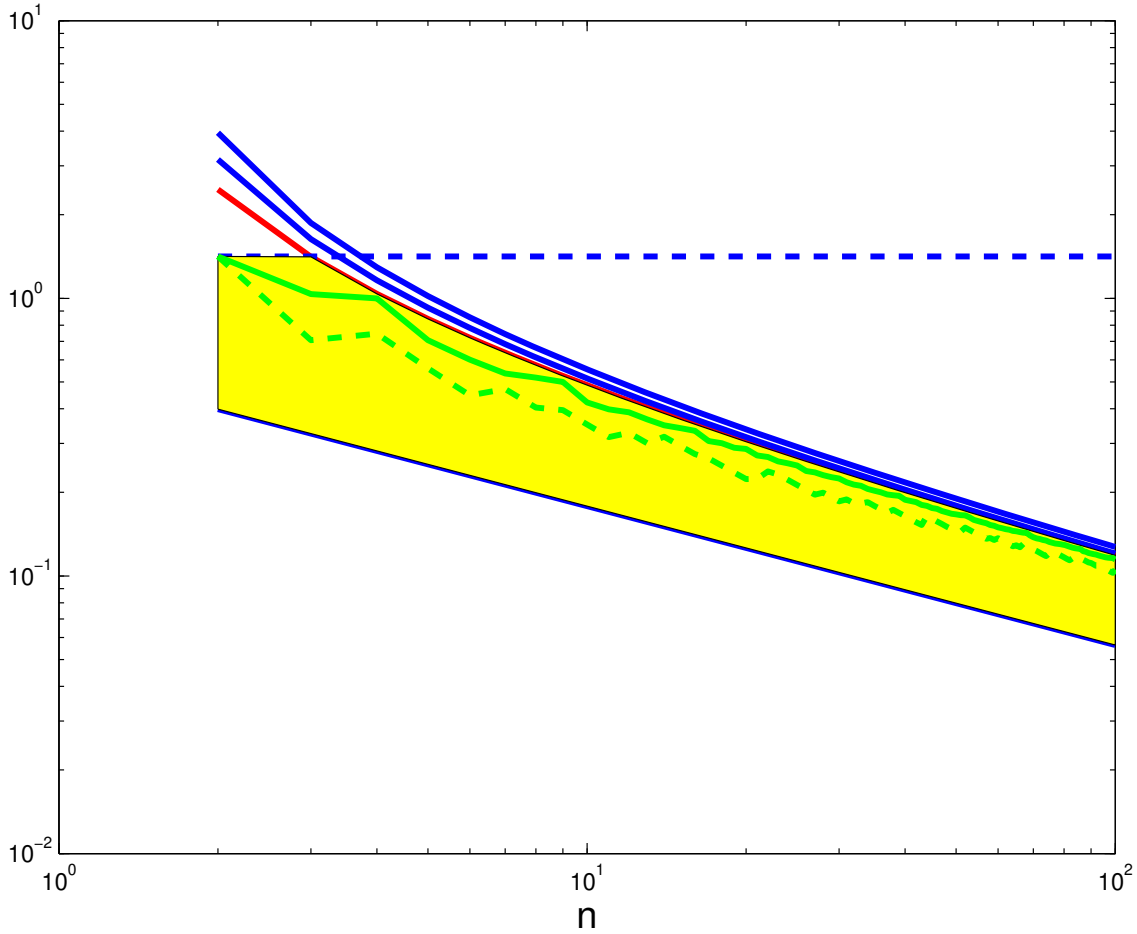
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 2$

$d = 2$, including $\Phi^*_{Mm,n}$ (proved up to $n = 30$ `http://www.packomania.com/`)



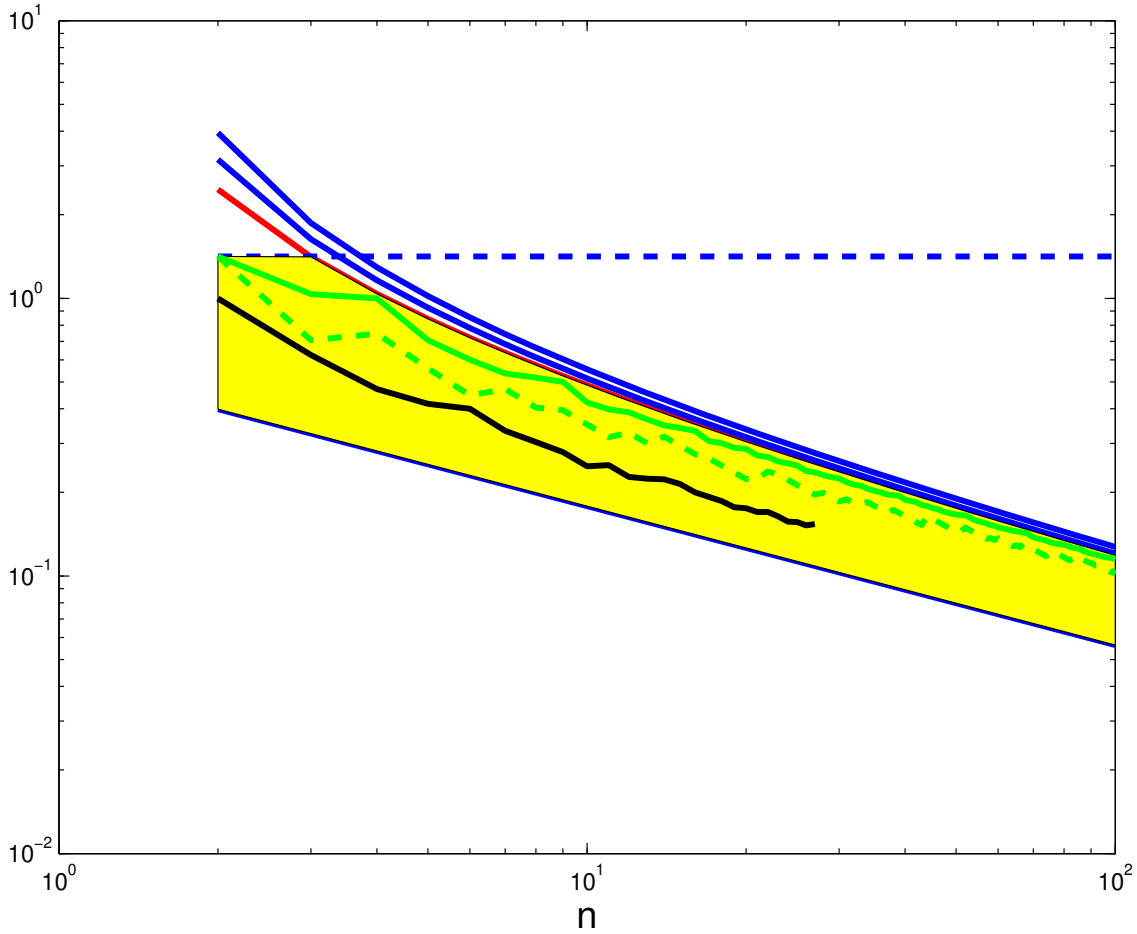
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 2$

$d = 2$, including $\Phi^*_{Mm-Lh,n}$ (proved up to $n = 70$ van Dam et al. (2007), `http://www.spacefillingdesigns.nl/`)



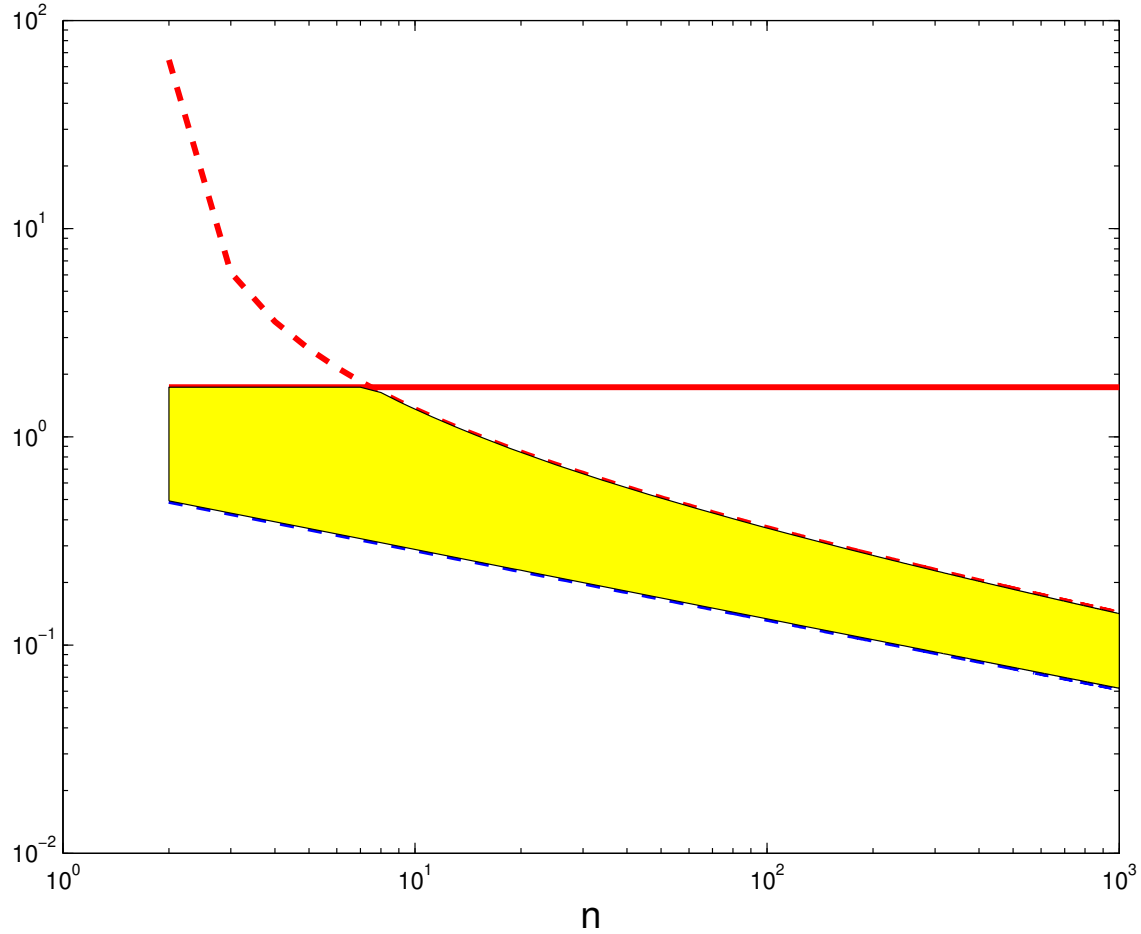
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 2$

$d = 2$, including $\Phi^*_{mM-Lh,n}$ (proved up to $n = 27$ (van Dam, 2008))



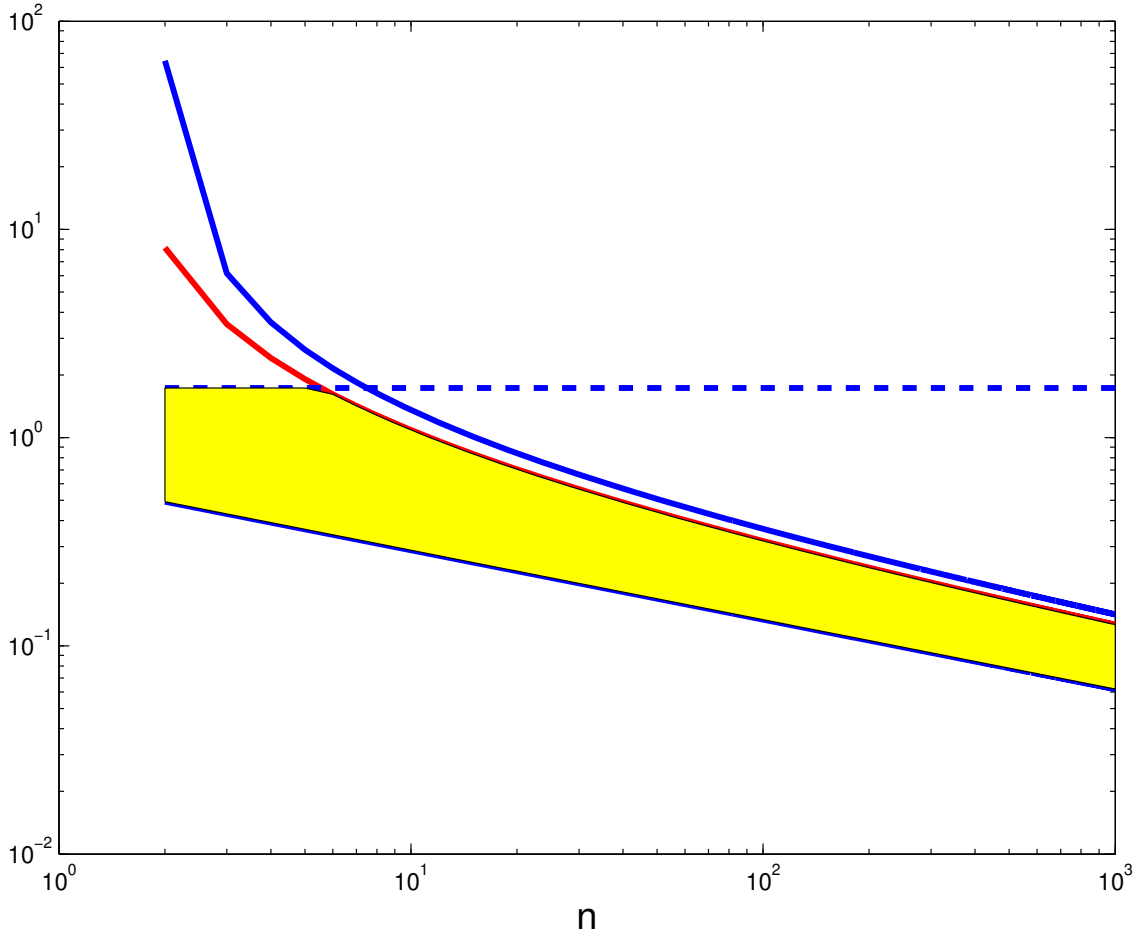
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 3$

$d = 3$ using $\boxed{R^*_n < \Phi^*_{mM,n} \leq \Phi^*_{Mm,n} < \min\{\frac{2R^*_n}{1-2R^*_n}, \sqrt{d}\}}$



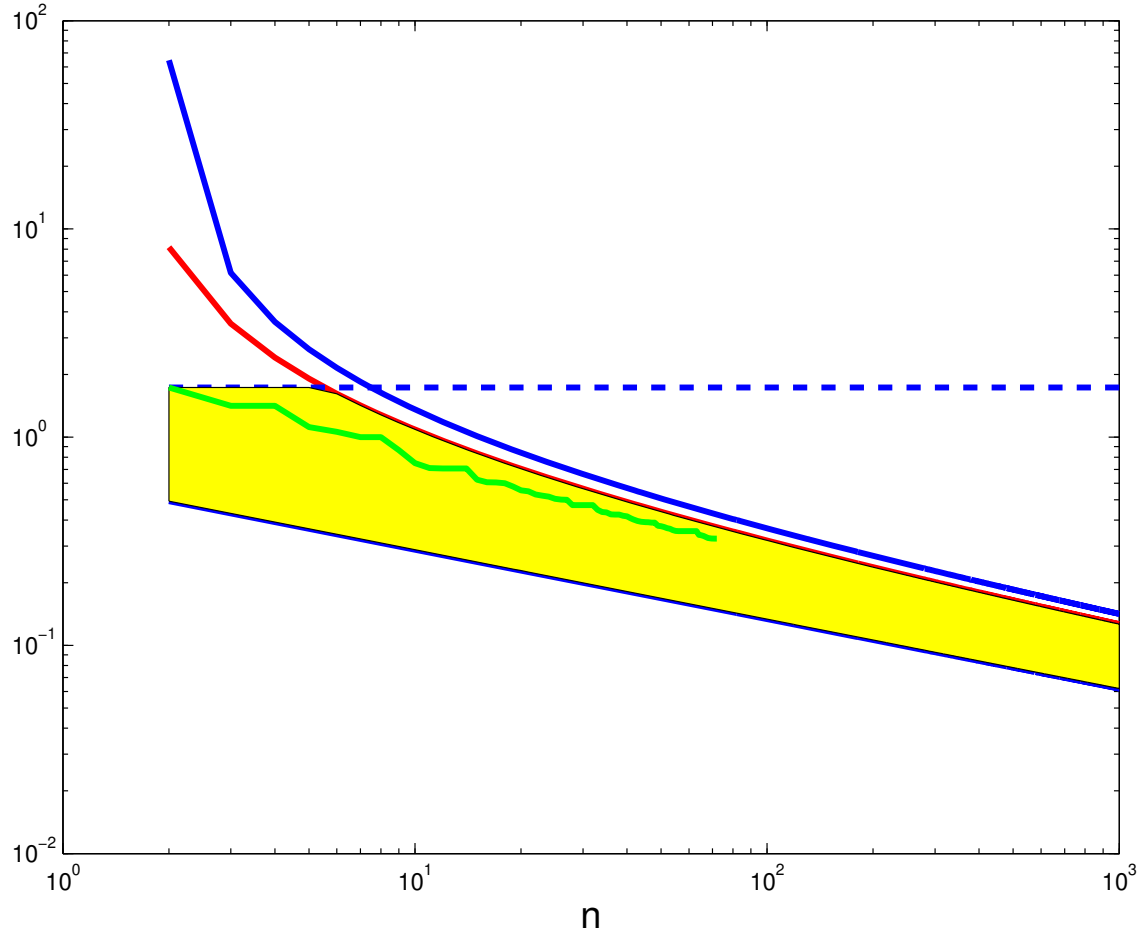
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 3$

$d = 3$ with moreover $\boxed{\Phi^*_{Mm,n} < \frac{1}{(n/\sqrt{2})^{1/3}-1}}$ (*packing density*)



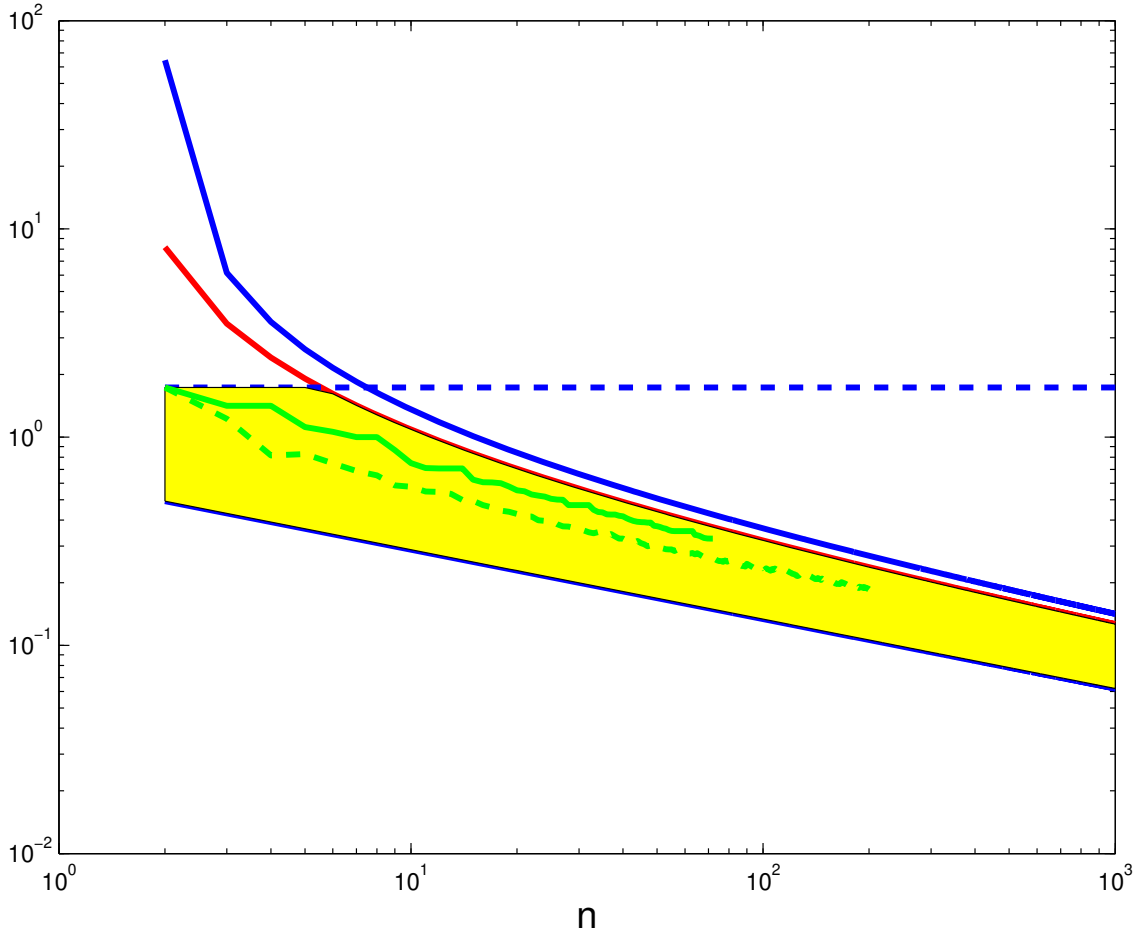
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 3$

$d = 3$, including $\Phi^*_{Mm,n}$ (`http://www.randomwalk.de/sphere/incube/`)



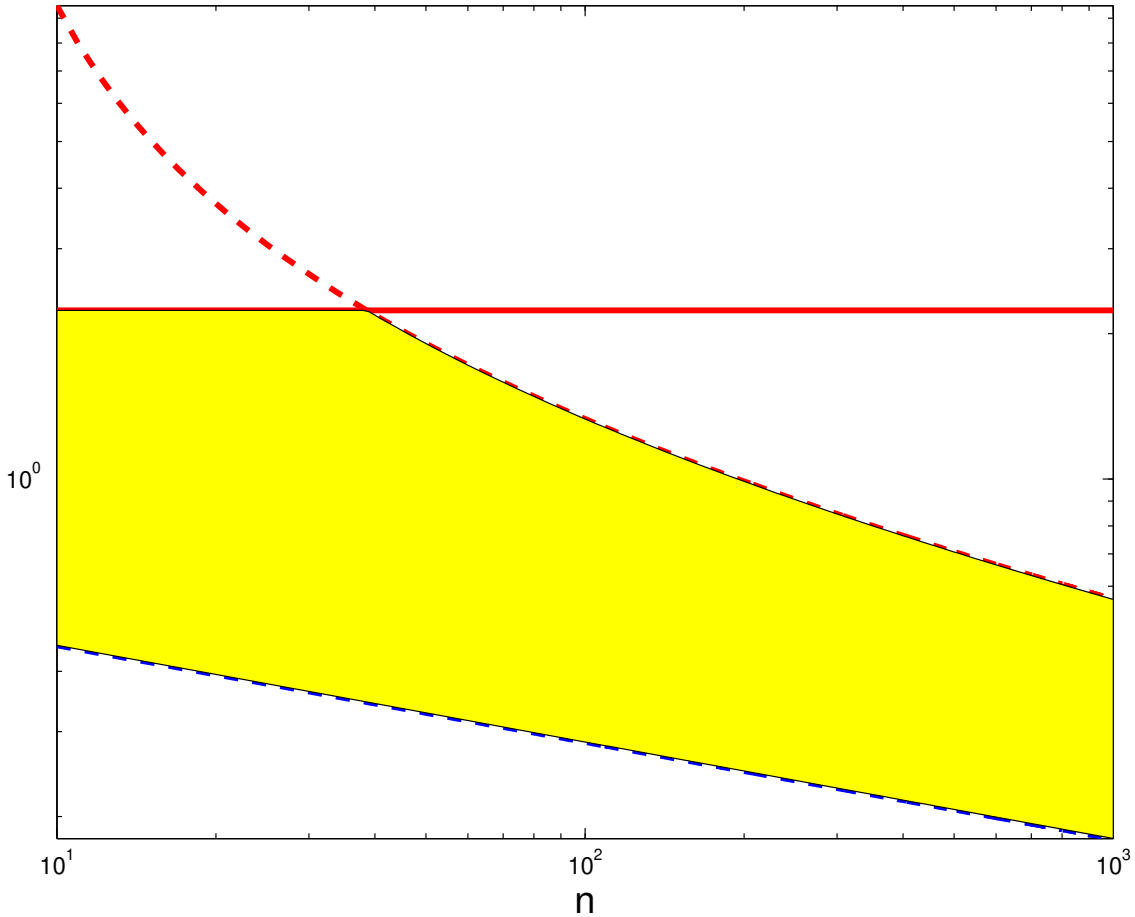
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 3$

$d = 3$, including $\Phi^*_{Mm-Lh,n}$ (proved up to $n = 15$ `http://www.spacefillingdesigns.nl/`)



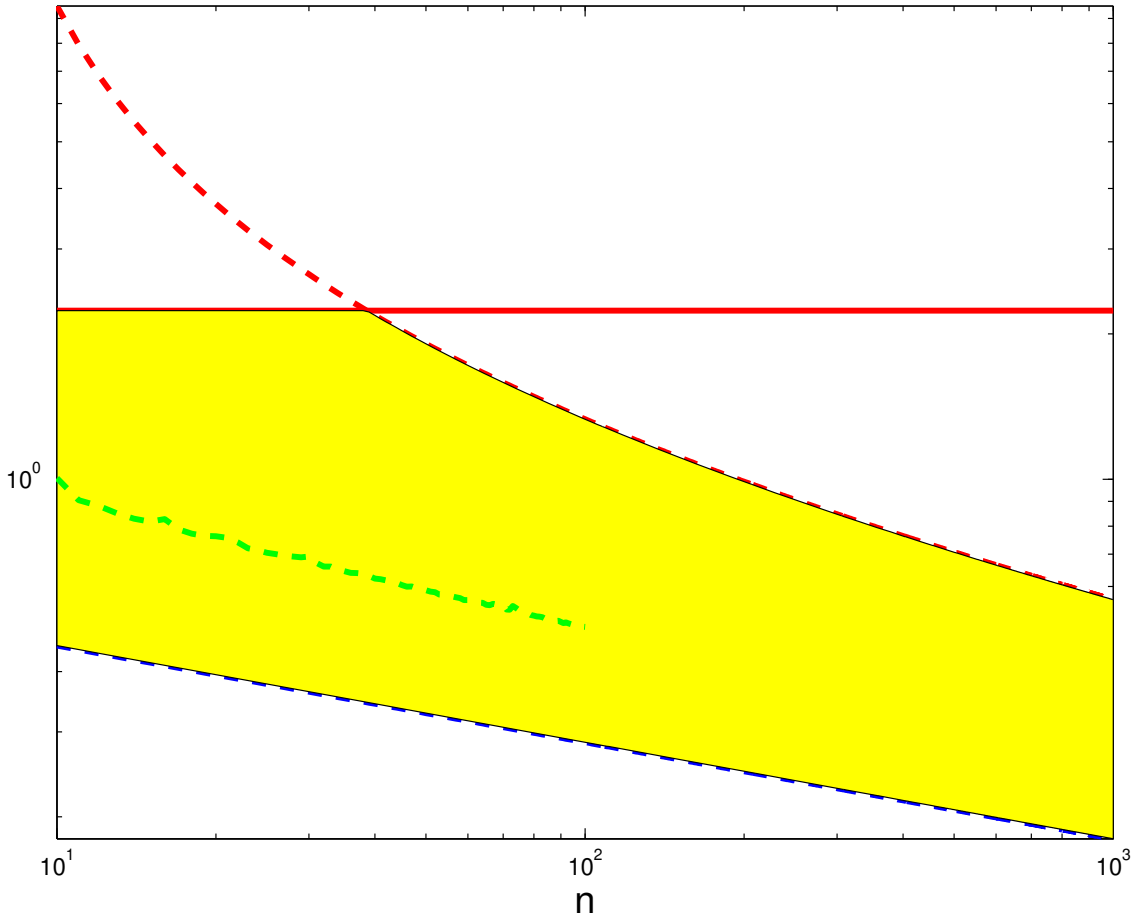
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 5$

$d = 5$ using $\boxed{R^*_n < \Phi^*_{mM,n} \leq \Phi^*_{Mm,n} < \min\{\frac{2R^*_n}{1-2R^*_n}, \sqrt{d}\}}$



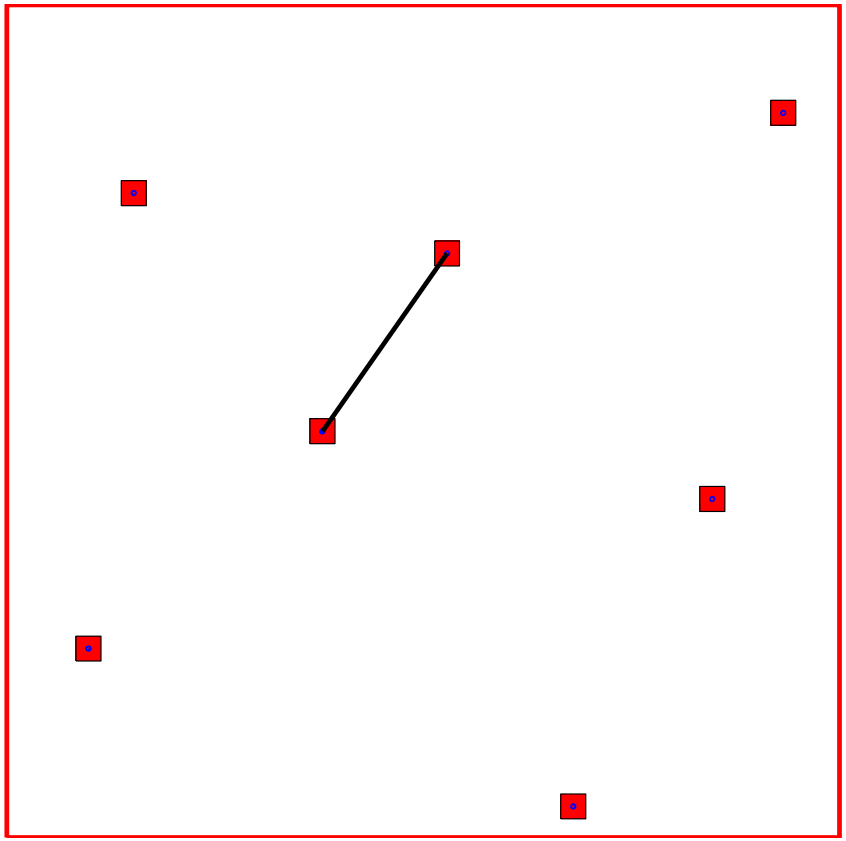
Bounds on $\Phi^*_{Mm,n}$ and $\Phi^*_{mM,n}$: $d = 5$

$d = 5$, including $\Phi^*_{Mm-Lh,n}$ (proved up to $n = 6$ http://www.spacefillingdesigns.nl/)



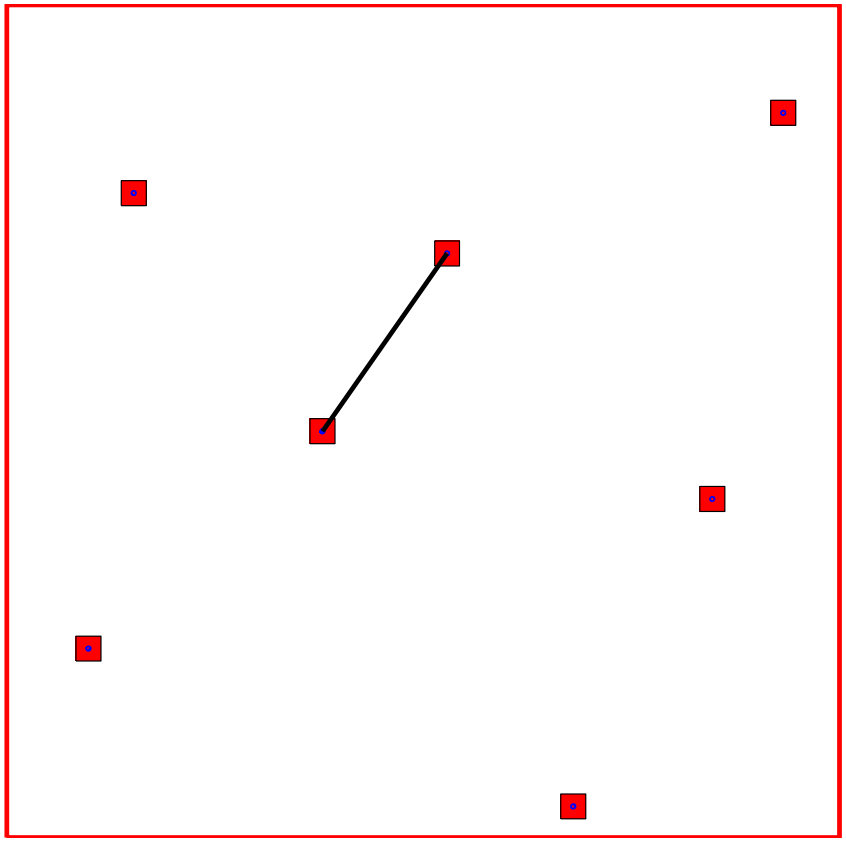
1.6 Regularized Maximin, energy

Maximin

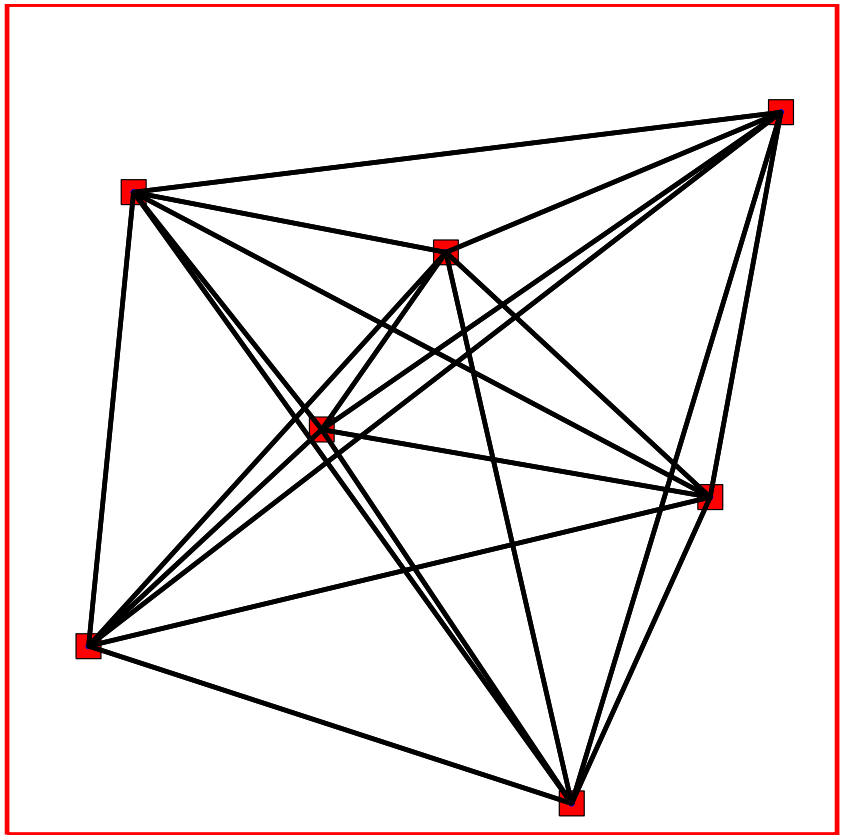


1.6 Regularized Maximin, energy

Maximin



Regularization: we account for distances between all pairs of points



Denote $d_{ij} \triangleq \|\mathbf{x}_i - \mathbf{x}_j\|$ ➠ $\Phi_{Mm}(\mathbf{X}_n) = \min_{i \neq j} d_{ij}$

$$\underline{\Phi}_{[q]}(\mathbf{X}_n) \triangleq \left[\sum_{i<j} d_{ij}^{-q} \right]^{-1/q} \quad \text{and} \quad \overline{\Phi}_{[q]}(\mathbf{X}_n) \triangleq \left[\frac{1}{N} \sum_{i<j} d_{ij}^{-q} \right]^{-1/q}$$

with $N = \binom{n}{2} = n(n-1)/2$

Denote $d_{ij} \triangleq \|\mathbf{x}_i - \mathbf{x}_j\|$ ➠ $\Phi_{Mm}(\mathbf{X}_n) = \min_{i \neq j} d_{ij}$

$$\underline{\Phi}_{[q]}(\mathbf{X}_n) \triangleq \left[\sum_{i<j} d_{ij}^{-q}\right]^{-1/q} \quad \text{and} \quad \overline{\Phi}_{[q]}(\mathbf{X}_n) \triangleq \left[\frac{1}{N}\sum_{i<j} d_{ij}^{-q}\right]^{-1/q}$$

with $N = \binom{n}{2} = n(n-1)/2$

Then, $\boxed{\underline{\Phi}_{[q]}(\mathbf{X}_n) \leq \Phi_{Mm}(\mathbf{X}_n) \leq \overline{\Phi}_{[q]}(\mathbf{X}_n) \leq N^{1/q}\,\underline{\Phi}_{[q]}(\mathbf{X}_n)}$, $q > 0$,
(monotonic convergence to $\Phi_{Mm}(\mathbf{X}_n)$ on both sides when $q \to \infty$)

Denote $d_{ij} \triangleq \|\mathbf{x}_i - \mathbf{x}_j\|$ ➠ $\Phi_{Mm}(\mathbf{X}_n) = \min_{i \neq j} d_{ij}$

$$\underline{\Phi}_{[q]}(\mathbf{X}_n) \triangleq \left[\sum_{i<j} d_{ij}^{-q}\right]^{-1/q} \quad \text{and} \quad \overline{\Phi}_{[q]}(\mathbf{X}_n) \triangleq \left[\frac{1}{N}\sum_{i<j} d_{ij}^{-q}\right]^{-1/q}$$

with $N = \binom{n}{2} = n(n-1)/2$

Then, $\boxed{\underline{\Phi}_{[q]}(\mathbf{X}_n) \leq \Phi_{Mm}(\mathbf{X}_n) \leq \overline{\Phi}_{[q]}(\mathbf{X}_n) \leq N^{1/q}\,\underline{\Phi}_{[q]}(\mathbf{X}_n)}$, $q > 0$,
(monotonic convergence to $\Phi_{Mm}(\mathbf{X}_n)$ on both sides when $q \to \infty$)

By continuity, $\overline{\Phi}_{[0]}(\mathbf{X}_n) = \exp\left[\frac{1}{N}\sum_{i<j}\log(d_{ij})\right]$

Let $\underline{\mathbf{X}}^*_{n[q]}$ be optimal for $\underline{\Phi}_{[q]}$:

$$\frac{\Phi_{Mm}(\underline{\mathbf{X}}^*_{n[q]})}{\Phi^*_{Mm,n}} \geq N^{-1/q}\,, \text{ tends to 1 as } q \to \infty$$

(Maximin efficiency $> 1 - \epsilon$ for $q > \frac{2\log(n)}{\epsilon}$)

Let $\underline{\mathbf{X}}^*_{n[q]}$ be optimal for $\underline{\Phi}_{[q]}$:

$$\frac{\Phi_{Mm}(\underline{\mathbf{X}}^*_{n[q]})}{\Phi^*_{Mm,n}} \geq N^{-1/q}\,, \text{ tends to 1 as } q \to \infty$$

(Maximin efficiency $> 1-\epsilon$ for $q > \frac{2\log(n)}{\epsilon}$)

➤ Maximize $\underline{\Phi}_{[q]}$ is equivalent to minimizing the energy
$E_q(\mathbf{X}_n) = \frac{2}{n(n-1)} \sum_{i,j=1,\, i\neq j}^{n} \|\mathbf{x}_i - \mathbf{x}_j\|^{-q}$

Audze and Eglais (1977) have proposed $q = 2$; when $q \lesssim 5$ optimization of Lh designs is easier than for Φ_{Mm} (Morris and Mitchell, 1995)

Let $\underline{\mathbf{X}}^*_{n[q]}$ be optimal for $\underline{\Phi}_{[q]}$:

$$\frac{\Phi_{Mm}(\underline{\mathbf{X}}^*_{n[q]})}{\Phi^*_{Mm,n}} \geq N^{-1/q}\,, \text{ tends to 1 as } q \to \infty$$

(Maximin efficiency $> 1-\epsilon$ for $q > \frac{2\log(n)}{\epsilon}$)

➤ Maximize $\underline{\Phi}_{[q]}$ is equivalent to minimizing the energy
$E_q(\mathbf{X}_n) = \frac{2}{n(n-1)} \sum_{i,j=1,\, i\neq j}^{n} \|\mathbf{x}_i - \mathbf{x}_j\|^{-q}$

Audze and Eglais (1977) have proposed $q = 2$; when $q \lesssim 5$ optimization of Lh designs is easier than for Φ_{Mm} (Morris and Mitchell, 1995)

Regularized version $\underline{\Phi}_{[q]}(\mathbf{X}_n)$: non-concave but differentiable
➠ local maximization "easy" for q not too large ...
but q should be large enough to get a good approximation of $\Phi_{Mm}(\mathbf{X}_n)$

Let $\underline{\mathbf{X}}^*_{n[q]}$ be optimal for $\underline{\Phi}_{[q]}$:

$$\frac{\Phi_{Mm}(\underline{\mathbf{X}}^*_{n[q]})}{\Phi^*_{Mm,n}} \geq N^{-1/q}\,, \text{ tends to 1 as } q \to \infty$$

(Maximin efficiency $> 1-\epsilon$ for $q > \frac{2\log(n)}{\epsilon}$)

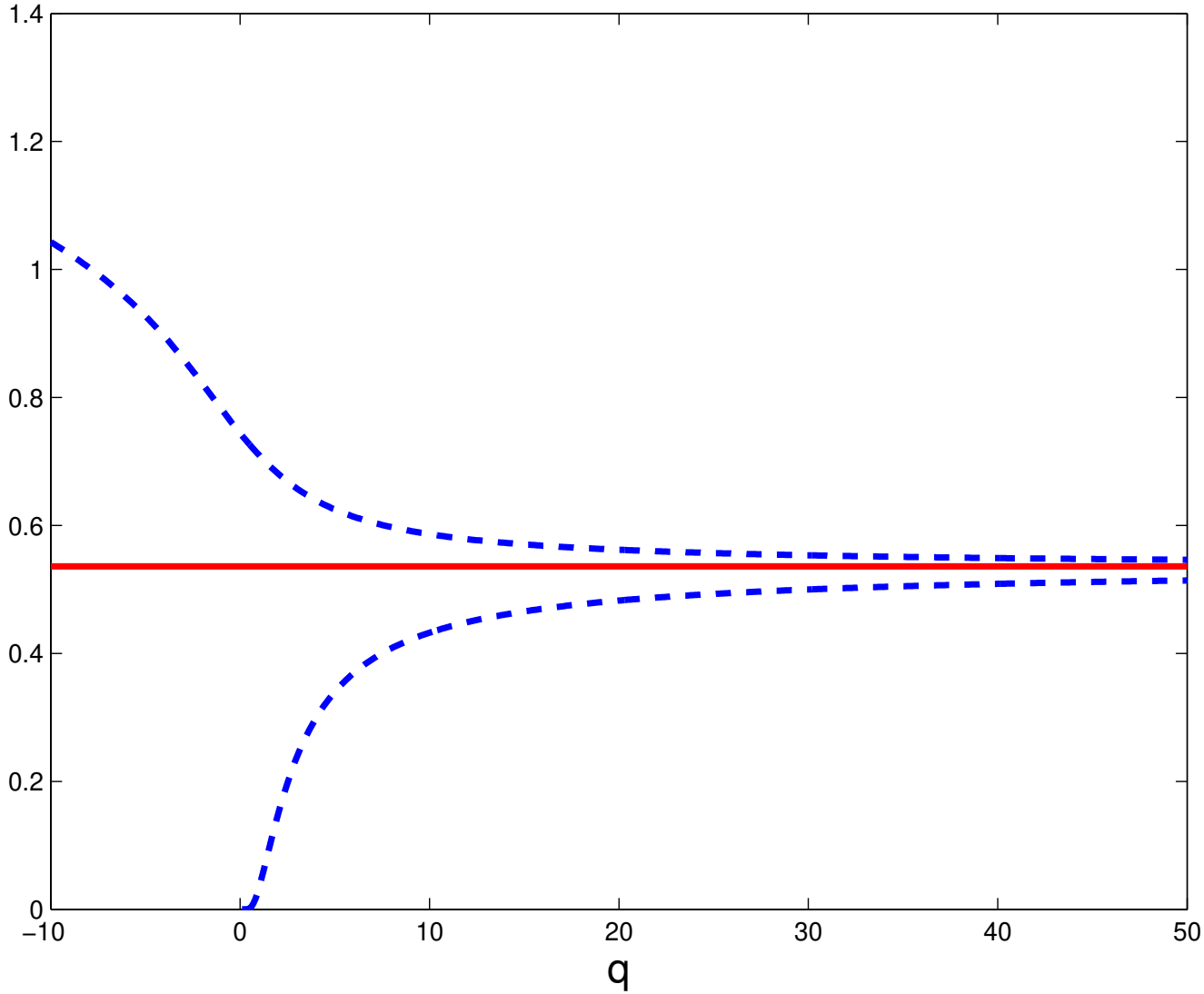
➤ Maximize $\underline{\Phi}_{[q]}$ is equivalent to minimizing the energy
$E_q(\mathbf{X}_n) = \frac{2}{n(n-1)} \sum_{i,j=1,\, i\neq j}^{n} \|\mathbf{x}_i - \mathbf{x}_j\|^{-q}$

Audze and Eglais (1977) have proposed $q = 2$; when $q \lesssim 5$ optimization of Lh designs is easier than for Φ_{Mm} (Morris and Mitchell, 1995)

Regularized version $\underline{\Phi}_{[q]}(\mathbf{X}_n)$: non-concave but differentiable
➠ local maximization "easy" for q not too large ...
but q should be large enough to get a good approximation of $\Phi_{Mm}(\mathbf{X}_n)$

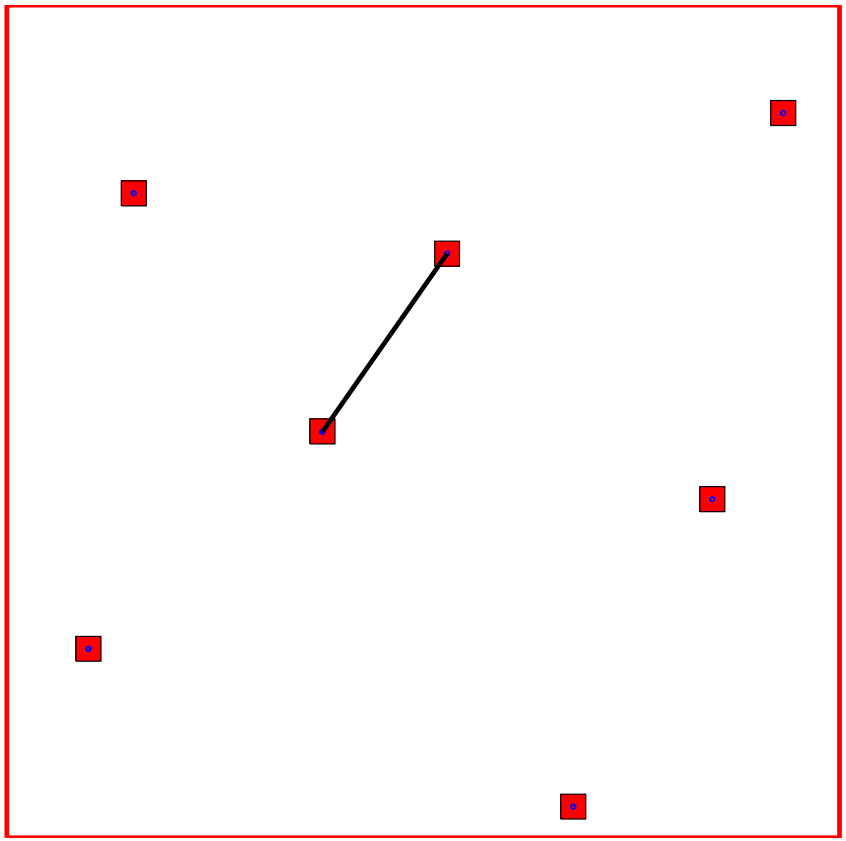
Relation with potential theory Landkof (1972); Saff (2010), P., Wynn and Zhigljavsky (2016): $\underline{\mathbf{X}}^*_{n[q]}$ = Fekete points, asymptotically distributed ($n \to \infty$) uniformly in $\mathscr{X}$ if $\boxed{q \geq d}$

$\mathbf{X}_n$ Maximin-optimal, $n = 7$, $d = 2$: Φ_{Mm} and bounds $\underline{\Phi}_{[q]}$ and $\overline{\Phi}_{[q]}$



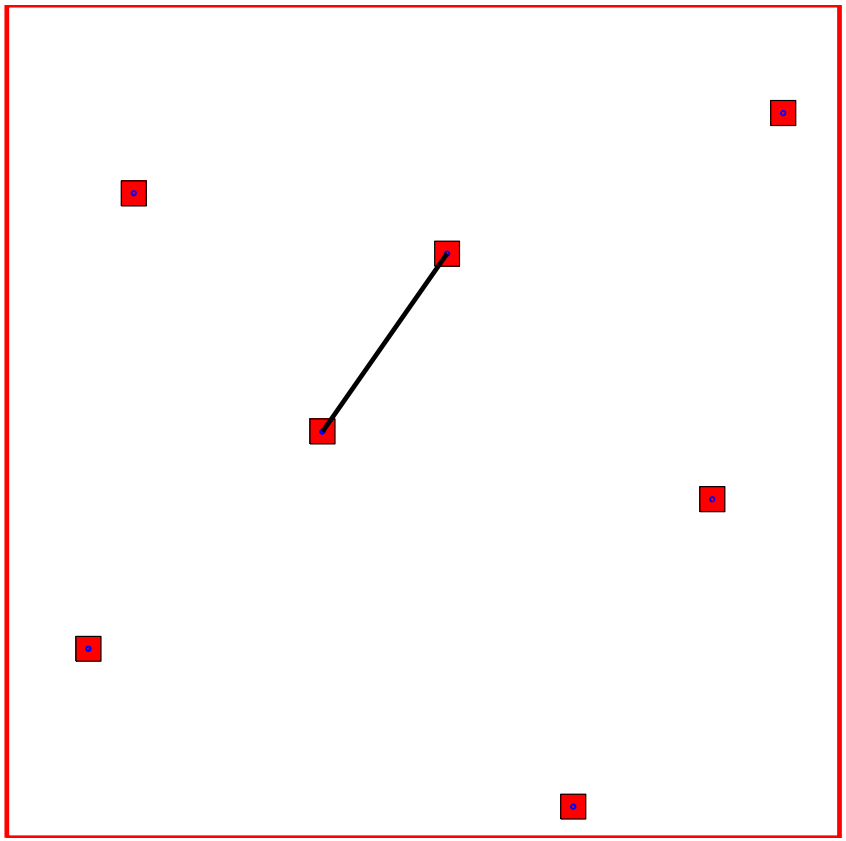
6) Regularized Maximin with Nearest Neighbors (NN)

Maximin

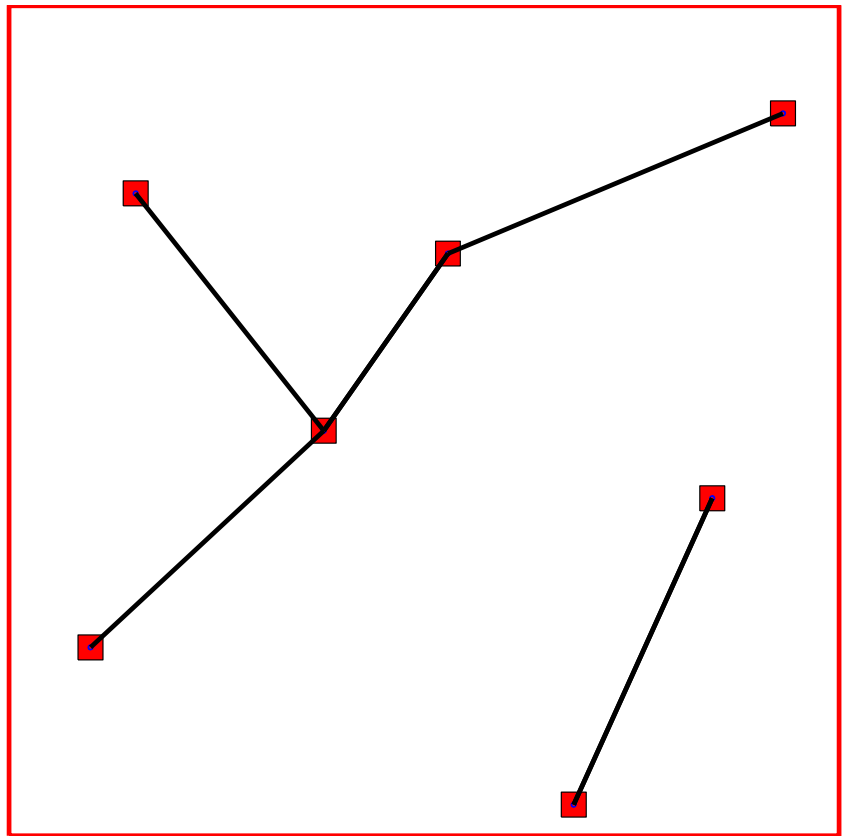


6) Regularized Maximin with Nearest Neighbors (NN)

Maximin



Regularization: we account for the distance between each point and its NN



We can write $\Phi_{Mm}(\mathbf{X}_n) = \min_i d_i^*$, with $d_i^* \triangleq \min_{j \neq i} \|\mathbf{x}_i - \mathbf{x}_j\|$
$=$ distance to NN of X_i

Define

$$\underline{\Phi}_{[NN,q]}(\mathbf{X}_n) = \left[\sum_{i=1}^{n} (d_i^*)^{-q}\right]^{-1/q} \quad , \quad \overline{\Phi}_{[NN,q]}(\mathbf{X}_n) = \left[\sum_{i=1}^{n} \frac{(d_i^*)^{-q}}{n}\right]^{-1/q}$$

(we only regularize $\min_i$)

We can write $\Phi_{Mm}(\mathbf{X}_n) = \min_i d_i^*$, with $d_i^* \triangleq \min_{j \neq i} \|\mathbf{x}_i - \mathbf{x}_j\|$
$=$ distance to NN of X_i

Define

$$\underline{\Phi}_{[NN,q]}(\mathbf{X}_n) = \left[\sum_{i=1}^{n} (d_i^*)^{-q}\right]^{-1/q} \quad , \quad \overline{\Phi}_{[NN,q]}(\mathbf{X}_n) = \left[\sum_{i=1}^{n} \frac{(d_i^*)^{-q}}{n}\right]^{-1/q}$$

(we only regularize $\min_i$)

Then, $\boxed{\underline{\Phi}_{[NN,q]}(\mathbf{X}_n) \leq \Phi_{Mm}(\mathbf{X}_n) \leq n^{1/q}\, \underline{\Phi}_{[NN,q]}(\mathbf{X}_n)}$, $q > 0$

(monotonic convergence to $\Phi_{Mm}(\mathbf{X}_n)$ on both sides when $q \to \infty$)

By continuity: $\overline{\Phi}_{[NN,0]}(\mathbf{X}_n) = \exp\left[\sum_{i=1}^{n} \frac{\log(d_i^*)}{n}\right]$

Maximin efficiency:

$$\frac{\Phi_{Mm}(\underline{\mathbf{X}}^*_{n[NN,q]})}{\Phi^*_{Mm,n}} \geq n^{-1/q},$$

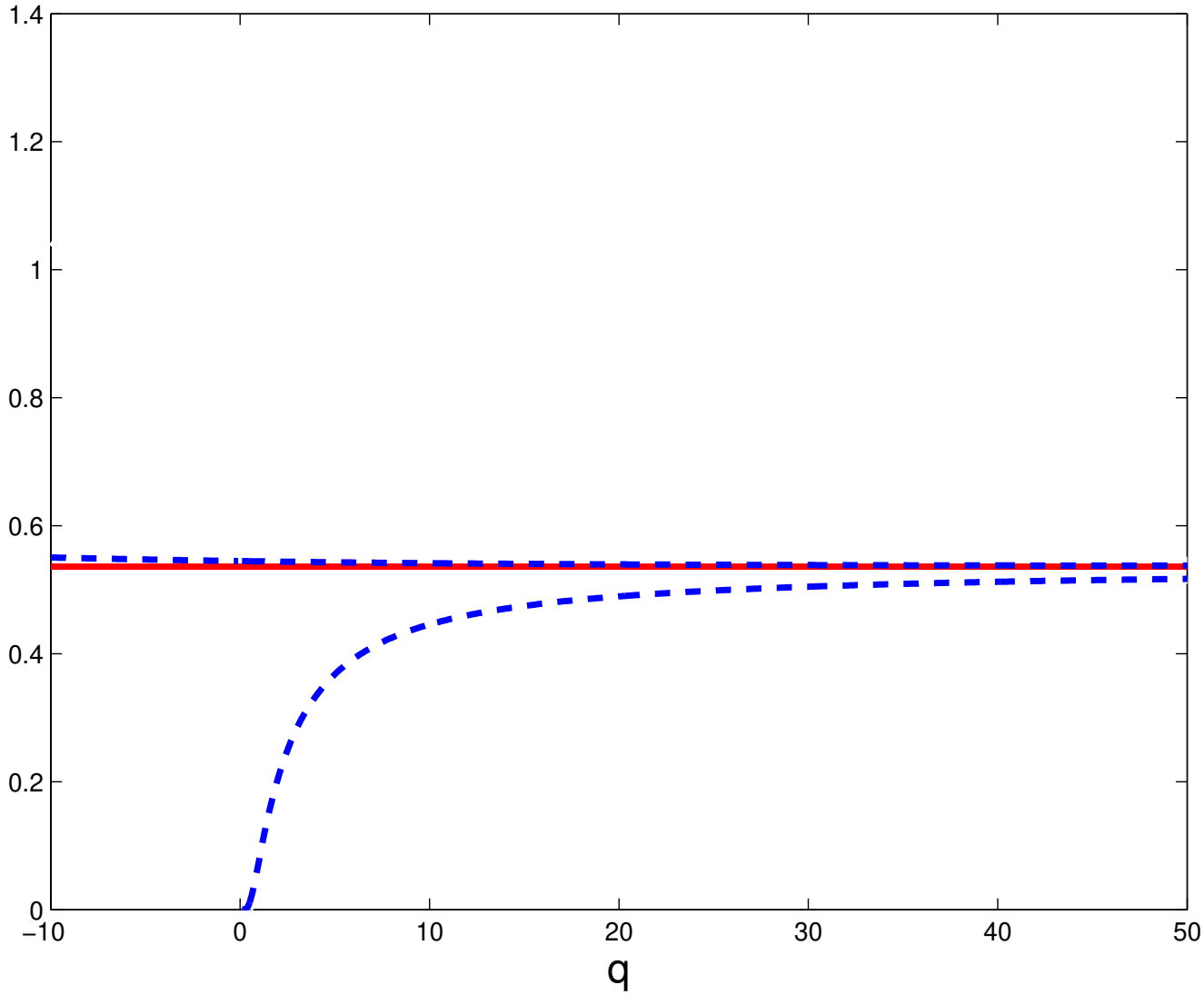
with $\underline{\mathbf{X}}^*_{n[NN,q]}$ optimal for $\underline{\Phi}_{[NN,q]}$

Maximin efficiency $> 1-\epsilon$ for $q > \frac{\log(n)}{\epsilon}$

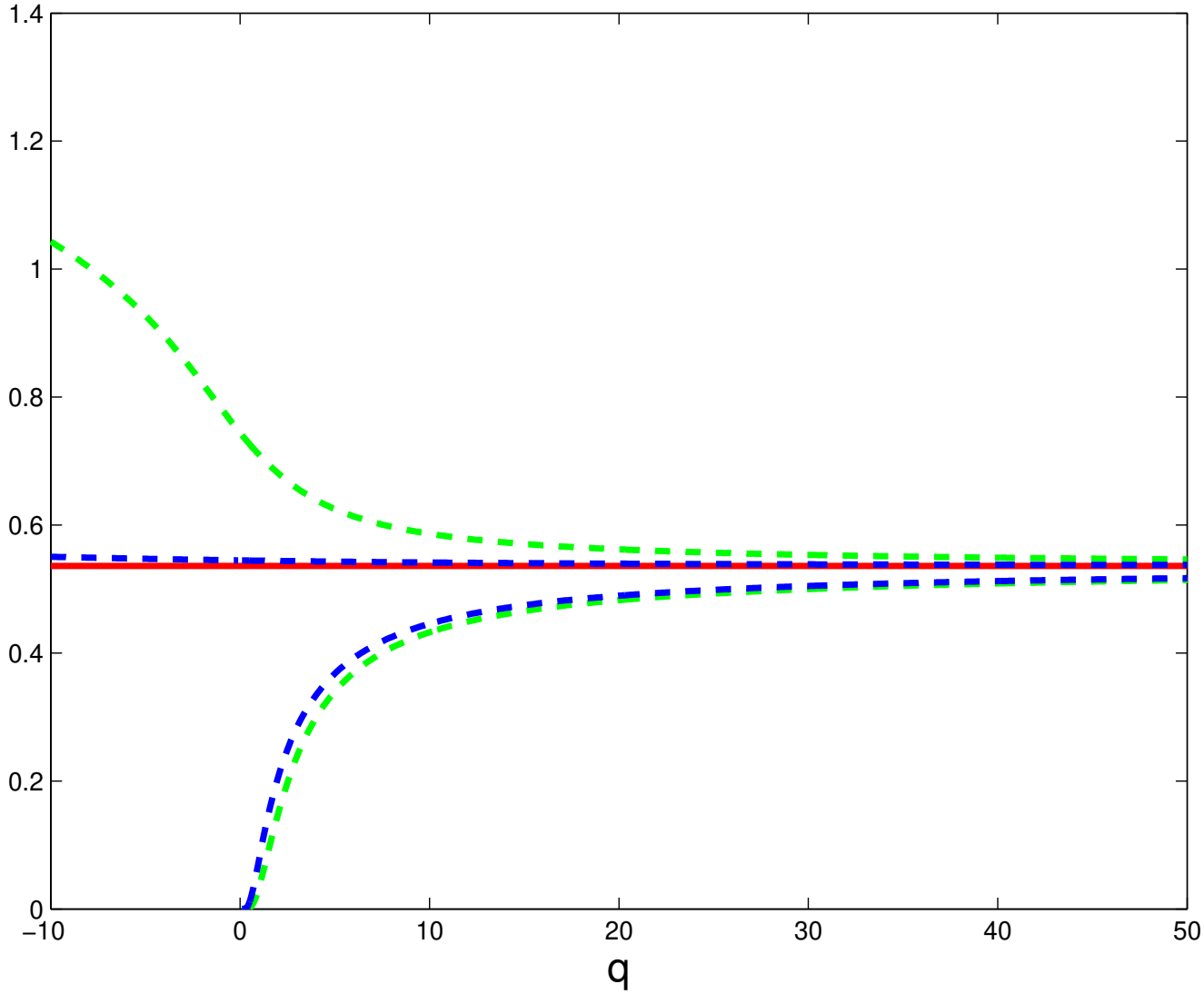
$\rightarrow$ we gain a factor 2 comparatively to $\underline{\Phi}_{[q]}(\mathbf{X}_n)$

(we only regularized $\min_i$)

$\mathbf{X}_n$ Maximin-optimal, $n = 7$, $d = 2$: Φ_{Mm} and regularization with $\underline{\Phi}_{[NN,q]}$



$\mathbf{X}_n$ Maximin-optimal, $n = 7$, $d = 2$: Φ_{Mm} and regularization with $\underline{\Phi}_{[NN,q]}$ and $\underline{\Phi}_{[q]}$



Maximum projection designs (Joseph et al., 2015)

➠ modification of regularized Maximin that produces designs with good space-filling properties in all lower dimensional subspaces

Maximum projection designs (Joseph et al., 2015)

➠ modification of regularized Maximin that produces designs with good space-filling properties in all lower dimensional subspaces

regularized Maximin: maximize $\underline{\Phi}_{[q]}(\mathbf{X}_n) \triangleq \left[\sum_{i<j} d_{ij}^{-q}\right]^{-1/q}$

where $d_{ij} \triangleq \|\mathbf{x}_i - \mathbf{x}_j\|$, $i, j = 1, \ldots, n$

$(\underline{\Phi}_{[q]}(\mathbf{X}_n) \to \Phi_{Mm}(\mathbf{X}_n) = \min_{i \neq j} d_{ij}$ as $q \to \infty)$

Replace ℓ_2 distance d_{ij} by weighted-ℓ_2 distance (measure of importance on factors)

$$d_{ij,\mathbf{w}} = \left[\sum_{k=1}^{d} w_k (\{\mathbf{x}_i\}_k - \{\mathbf{x}_j\}_k)^2\right]^{1/2}$$

➞ minimize $\sum_{i<j} d_{ij,\mathbf{w}}^{-q}$ for a large q, but which $\mathbf{w}$?

Maximum projection designs (Joseph et al., 2015)

➠ modification of regularized Maximin that produces designs with good space-filling properties in all lower dimensional subspaces

regularized Maximin: maximize $\underline{\Phi}_{[q]}(\mathbf{X}_n) \triangleq \left[\sum_{i<j} d_{ij}^{-q}\right]^{-1/q}$

where $d_{ij} \triangleq \|\mathbf{x}_i - \mathbf{x}_j\|$, $i, j = 1, \ldots, n$

$(\underline{\Phi}_{[q]}(\mathbf{X}_n) \to \Phi_{Mm}(\mathbf{X}_n) = \min_{i \neq j} d_{ij}$ as $q \to \infty)$

Replace ℓ_2 distance d_{ij} by weighted-ℓ_2 distance (measure of importance on factors)

$$d_{ij,\mathbf{w}} = \left[\sum_{k=1}^{d} w_k (\{\mathbf{x}_i\}_k - \{\mathbf{x}_j\}_k)^2\right]^{1/2}$$

➞ minimize $\sum_{i<j} d_{ij,\mathbf{w}}^{-q}$ for a large q, but which $\mathbf{w}$?

Take $\mathbf{w} \in \mathcal{P}_d = \{\mathbf{w} : w_k \geq 0,\ \sum_{k=1}^{d} w_k = 1\}$

Put a uniform prior π on $(w_1, \ldots, w_{d-1})$

For $q = 2d$, $\boxed{\mathrm{E}_{\mathbf{w}}\left\{\sum_{i<j} d_{ij,\mathbf{w}}^{-q}\right\} = \frac{1}{[(d-1)!]^2} \sum_{i<j} \frac{1}{\prod_{k=1}^{d} (\{\mathbf{x}_i\}_k - \{\mathbf{x}_j\}_k)^2}}$ Very promising!

Remark: measures of regularity (and not of space fillingness!)

$$R_1(\mathbf{X}_n) \triangleq \frac{\text{var}^{1/2}(d_i^*)}{\text{E}(d_i^*)} = \frac{\left[\frac{1}{n}\sum_{i=1}^{n}(d_i^* - \bar{d})^2\right]^{1/2}}{\bar{d}}, \text{ with } \bar{d} \triangleq \text{E}(d_i^*) = \frac{1}{n}\sum_{i=1}^{n} d_i^*$$

$= 0$ for a regular grid

▲ invariant by scale transformation ▲

sometimes called "covering measure"... which it is not

Remark: measures of regularity (and not of space fillingness!)

$R_1(\mathbf{X}_n) \triangleq \frac{\text{var}^{1/2}(d_i^*)}{\text{E}(d_i^*)} = \frac{\left[\frac{1}{n}\sum_{i=1}^{n}(d_i^* - \bar{d})^2\right]^{1/2}}{\bar{d}}$, with $\bar{d} \triangleq \text{E}(d_i^*) = \frac{1}{n}\sum_{i=1}^{n} d_i^*$

$= 0$ for a regular grid

▲ invariant by scale transformation ▲

sometimes called "covering measure"... which it is not

$R_2(\mathbf{X}_n) \triangleq \frac{\max_{1\leq i\leq n} d_i^*}{\min_{1\leq i\leq n} d_i^*}$ (≥ 1, $= 1$ for a regular grid)

▲ invariant by scale transformation ▲

Remark: measures of regularity (and not of space fillingness!)

$$R_1(\mathbf{X}_n) \triangleq \frac{\text{var}^{1/2}(d_i^*)}{\text{E}(d_i^*)} = \frac{\left[\frac{1}{n}\sum_{i=1}^{n}(d_i^* - \bar{d})^2\right]^{1/2}}{\bar{d}}, \text{ with } \bar{d} \triangleq \text{E}(d_i^*) = \frac{1}{n}\sum_{i=1}^{n} d_i^*$$

$= 0$ for a regular grid

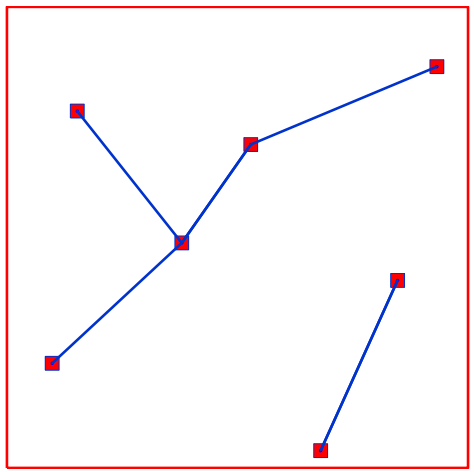
▲ invariant by scale transformation ▲

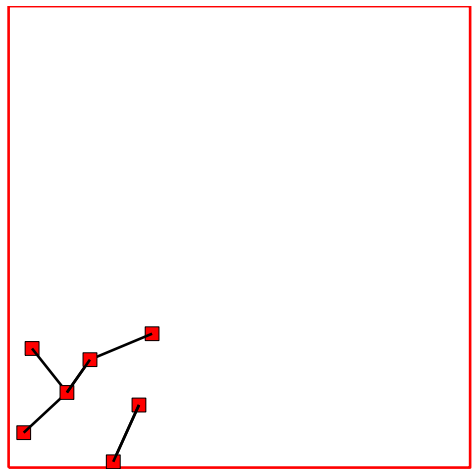
sometimes called "covering measure"... which it is not

$$R_2(\mathbf{X}_n) \triangleq \frac{\max_{1\leq i\leq n} d_i^*}{\min_{1\leq i\leq n} d_i^*} \quad (\geq 1, = 1 \text{ for a regular grid})$$

▲ invariant by scale transformation ▲

These two designs have the same R_1 and R_2 values





2 Uniformity: quasi Monte-Carlo, discrepancy

2.1 Entropy, optimal graphs

Consider $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ as a sample of size n of variables $\mathbf{x}_i$ i.i.d. in $\mathscr{X}$ with p.d.f. $\varphi(\cdot)$

Rényi entropy of $\varphi(\cdot)$ of order α :

$$H_\alpha{}^*(\varphi) \triangleq \frac{1}{1-\alpha} \log \int \varphi^\alpha(\mathbf{x})\,\mathrm{d}\mathbf{x} \ (\alpha \neq 1)$$

Tsallis entropy of $\varphi(\cdot)$ of order α :

$$H_\alpha(\varphi) \triangleq \frac{1}{\alpha-1}\left[1 - \int \varphi^\alpha(\mathbf{x})\,\mathrm{d}\mathbf{x}\right] \ (\alpha \neq 1)$$

which tend to

$$H_1(\varphi) \triangleq -\log \int \varphi(\mathbf{x}) \log[\varphi(\mathbf{x})]\,\mathrm{d}\mathbf{x} \text{ (Shannon entropy) as } \alpha \to 1$$

2 Uniformity: quasi Monte-Carlo, discrepancy

2.1 Entropy, optimal graphs

Consider $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ as a sample of size n of variables $\mathbf{x}_i$ i.i.d. in $\mathscr{X}$ with p.d.f. $\varphi(\cdot)$

Rényi entropy of $\varphi(\cdot)$ of order α :

$$H_\alpha{}^*(\varphi) \triangleq \frac{1}{1-\alpha} \log \int \varphi^\alpha(\mathbf{x})\, \mathrm{d}\mathbf{x} \ (\alpha \neq 1)$$

Tsallis entropy of $\varphi(\cdot)$ of order α :

$$H_\alpha(\varphi) \triangleq \frac{1}{\alpha-1} \left[1 - \int \varphi^\alpha(\mathbf{x})\, \mathrm{d}\mathbf{x}\right] \ (\alpha \neq 1)$$

which tend to

$$H_1(\varphi) \triangleq -\log \int \varphi(\mathbf{x}) \log[\varphi(\mathbf{x})]\, \mathrm{d}\mathbf{x} \text{ (Shannon entropy) as } \alpha \to 1$$

➤ For $\alpha > 0$, $H_\alpha{}^*(\varphi)$, $H_\alpha(\varphi)$ maximum for φ uniform over $\mathscr{X}$

2 Uniformity: quasi Monte-Carlo, discrepancy

2.1 Entropy, optimal graphs

Consider $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ as a sample of size n of variables $\mathbf{x}_i$ i.i.d. in $\mathscr{X}$ with p.d.f. $\varphi(\cdot)$

Rényi entropy of $\varphi(\cdot)$ of order α :

$$H_\alpha{}^*(\varphi) \triangleq \frac{1}{1-\alpha} \log \int \varphi^\alpha(\mathbf{x})\, \mathrm{d}\mathbf{x} \ (\alpha \neq 1)$$

Tsallis entropy of $\varphi(\cdot)$ of order α :

$$H_\alpha(\varphi) \triangleq \frac{1}{\alpha-1} \left[1 - \int \varphi^\alpha(\mathbf{x})\, \mathrm{d}\mathbf{x}\right] \ (\alpha \neq 1)$$

which tend to

$$H_1(\varphi) \triangleq -\log \int \varphi(\mathbf{x}) \log[\varphi(\mathbf{x})]\, \mathrm{d}\mathbf{x} \text{ (Shannon entropy) as } \alpha \to 1$$

➤ For $\alpha > 0$, $H_\alpha{}^*(\varphi)$, $H_\alpha(\varphi)$ maximum for φ uniform over $\mathscr{X}$

➠ Construct an estimate $\hat{H}_{n\alpha}$ of $H_\alpha^*(\varphi)$ from $\mathbf{X}_n$, use $\hat{H}_{n\alpha}$ as design criterion, to be maximized w.r.t. $\mathbf{X}_n$

▲ Entropie of "distribution of $\mathbf{x}_i$" $\neq$ entropy criterion for Gaussian Random Fields, see § II-1.3 (although relations exist)

▲ Entropie of "distribution of $\mathbf{x}_i$" $\neq$ entropy criterion for Gaussian Random Fields, see § II-1.3 (although relations exist)

1) A rather natural idea: plug-in

- Construct a kernel estimator $\hat{\varphi}_n(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} K_{\sigma^2}(\mathbf{x} - \mathbf{x}_i)$, $K_{\sigma^2}(\cdot)$ = p.d.f. with mean $\mathbf{0}$ and variance $\sigma^2\mathbf{I}_d$ (small enough)
- Use $H_\alpha(\hat{\varphi}_n)$, $\alpha > 0$, as design criterion

Jourdan and Franco (2010) use $H_1(\varphi)$, computationally costly for large d

▲ Entropie of "distribution of $\mathbf{x}_i$" $\neq$ entropy criterion for Gaussian Random Fields, see § II-1.3 (although relations exist)

1) A rather natural idea: plug-in

- Construct a kernel estimator $\hat{\varphi}_n(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^n K_{\sigma^2}(\mathbf{x}-\mathbf{x}_i)$, $K_{\sigma^2}(\cdot)$ = p.d.f. with mean $\mathbf{0}$ and variance $\sigma^2\mathbf{I}_d$ (small enough)
- Use $H_\alpha(\hat{\varphi}_n)$, $\alpha > 0$, as design criterion

Jourdan and Franco (2010) use $H_1(\varphi)$, computationally costly for large d

A peculiarity of H_2: if $K_{\sigma^2}(\cdot)$ corresponds to $\mathcal{N}(0,\sigma^2)$, then

$$\int_{\mathbb{R}^d} \hat{\varphi}_n^2(\mathbf{x})\,\mathrm{d}\mathbf{x} = \frac{1}{n^2}\sum_{i,j=1}^n K_{2\sigma^2}(\mathbf{x}_i-\mathbf{x}_j)$$

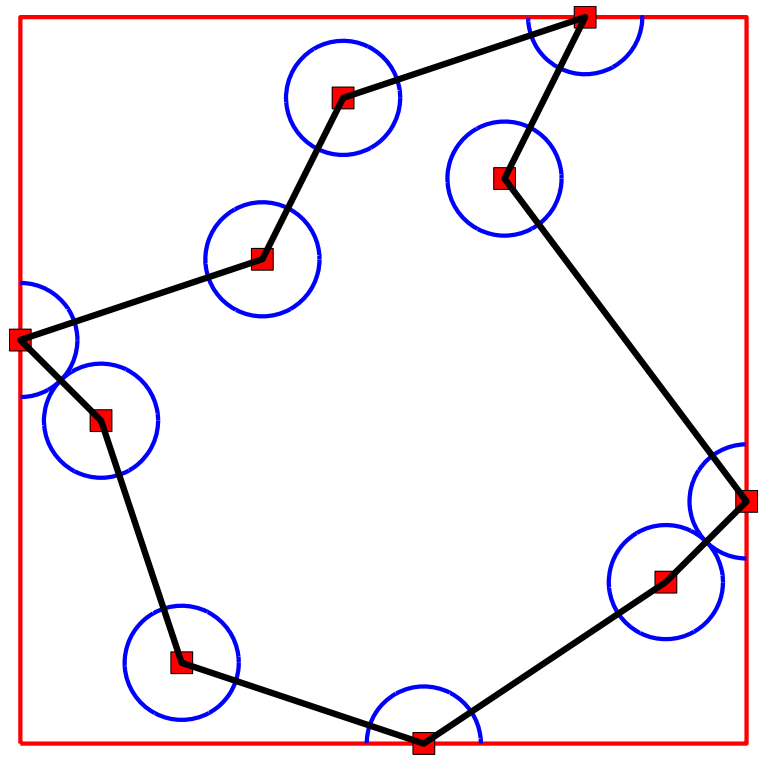
$$\Rightarrow \boxed{H_2(\hat{\varphi}_n) = 1 - \frac{1}{n^2}\sum_{i,j=1}^n K_{2\sigma^2}(\mathbf{x}_i-\mathbf{x}_j)}$$

= intra-distances criterion

2) Optimal graphs:

n points $\mathbf{X}_n$
(here, a Lh with $n = 10$, $d = 2$)

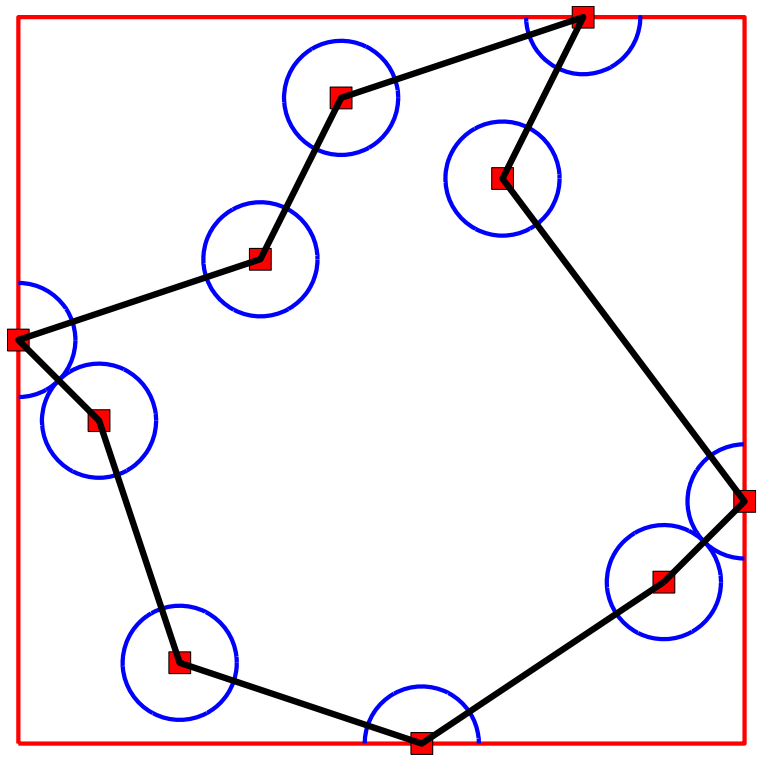
Traveling Salesman (TS) graph $\mathcal{G}_{TS}(\mathbf{X}_n)$



2) Optimal graphs:

n points $\mathbf{X}_n$
(here, a Lh with $n = 10$, $d = 2$)

Traveling Salesman (TS) graph $\mathcal{G}_{TS}(\mathbf{X}_n)$



(Beardwood et al., 1959): $\mathbf{x}_i$ i.i.d. with p.d.f. φ, the edges e_i of $\mathcal{G}_{TS}(\mathbf{X}_n)$ satisfy:

$$\frac{\sum_{e_i \in \mathcal{G}_{TS}(\mathbf{X}_n)} |e_i|}{n^{(d-1)/d}} \to C(d) \int \varphi^{(d-1)/d}(\mathbf{x})\, \mathrm{d}\mathbf{x} \text{ a.s.}, \quad n \to \infty$$

Later (Steele, 1981) considered other Euclidean functionals on $\mathbf{X}_n$, (Redmond and Yukich, 1994) used the notion of quasi-additivity

Many results... (Redmond and Yukich, 1996; Yukich, 1998; Penrose and Yukich, 2003; Wade, 2007; Penrose and Yukich, 2011)...

$$\frac{\sum_{e_i \in \mathcal{G}(\mathbf{X}_n)} |e_i|^\beta}{n^{1-\beta/d}} \to C(\beta, d) \int \varphi^{1-\beta/d}(\mathbf{x})\, \mathrm{d}\mathbf{x}\,, \quad n \to \infty$$

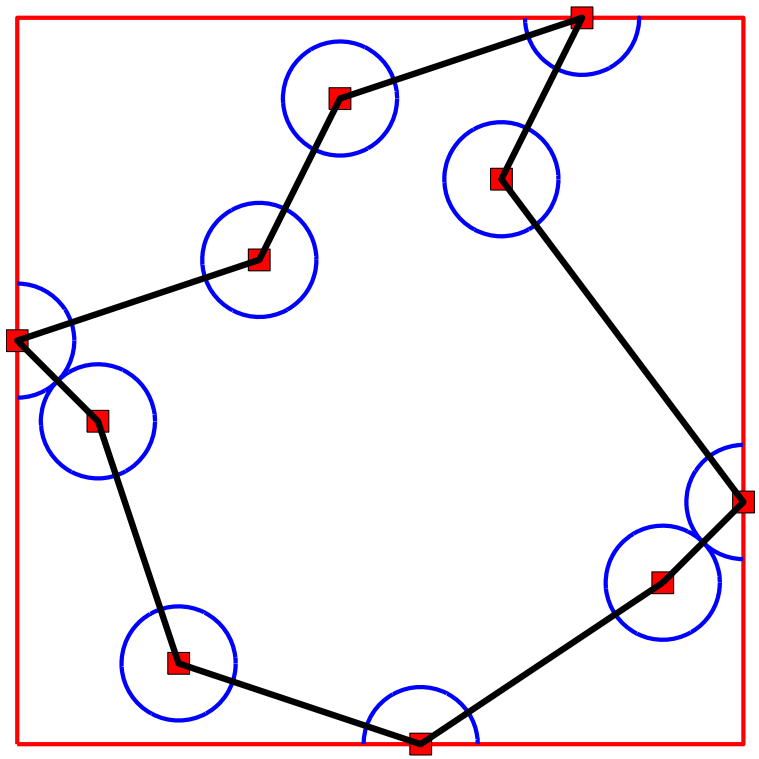
with $\mathcal{G}(\mathbf{X}_n)$ Minimum Spanning Tree (MST), NN, TS, Voronoï, Delaunay, Sphere of Influence, Gabriel... (different types of convergence (L_p), different conditions on φ and β...)

Many results... (Redmond and Yukich, 1996; Yukich, 1998; Penrose and Yukich, 2003; Wade, 2007; Penrose and Yukich, 2011)...

$$\frac{\sum_{e_i \in \mathcal{G}(\mathbf{X}_n)} |e_i|^\beta}{n^{1-\beta/d}} \to C(\beta, d) \int \varphi^{1-\beta/d}(\mathbf{x})\, \mathrm{d}\mathbf{x}\,, \quad n \to \infty$$

with $\mathcal{G}(\mathbf{X}_n)$ Minimum Spanning Tree (MST), NN, TS, Voronoï, Delaunay, Sphere of Influence, Gabriel... (different types of convergence (L_p), different conditions on φ and β...)

$\mathbf{X}_n$ a Lh with $n = 10$, $d = 2$: $\mathcal{G}_{TS}(\mathbf{X}_n)$

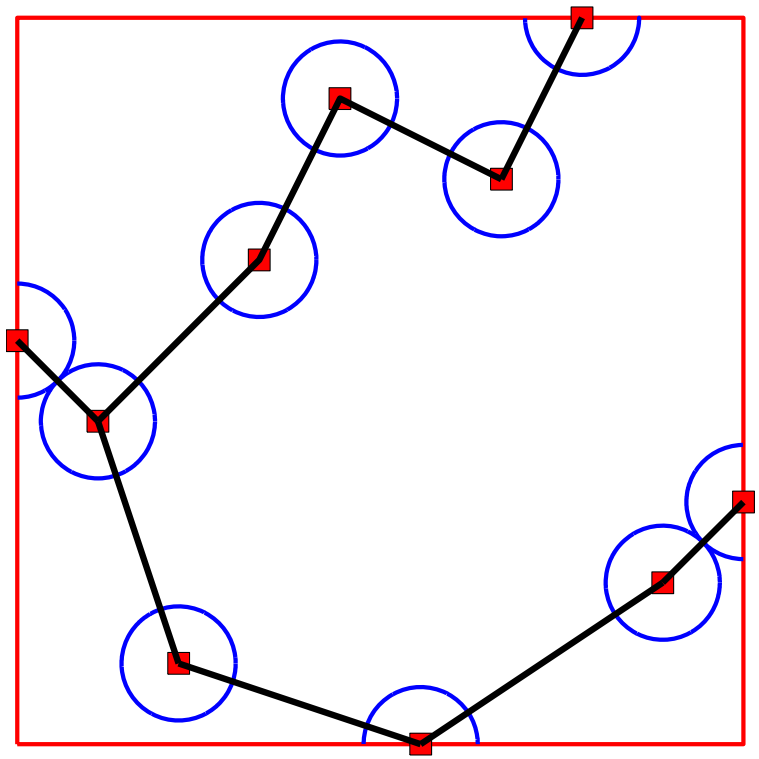


Many results... (Redmond and Yukich, 1996; Yukich, 1998; Penrose and Yukich, 2003; Wade, 2007; Penrose and Yukich, 2011)...

$$\frac{\sum_{e_i \in \mathcal{G}(\mathbf{X}_n)} |e_i|^\beta}{n^{1-\beta/d}} \to C(\beta, d) \int \varphi^{1-\beta/d}(\mathbf{x})\, \mathrm{d}\mathbf{x}\,, \quad n \to \infty$$

with $\mathcal{G}(\mathbf{X}_n)$ Minimum Spanning Tree (MST), NN, TS, Voronoï, Delaunay, Sphere of Influence, Gabriel... (different types of convergence (L_p), different conditions on φ and β...)

$\mathbf{X}_n$ a Lh with $n = 10$, $d = 2$: $\mathcal{G}_{MST}(\mathbf{X}_n)$

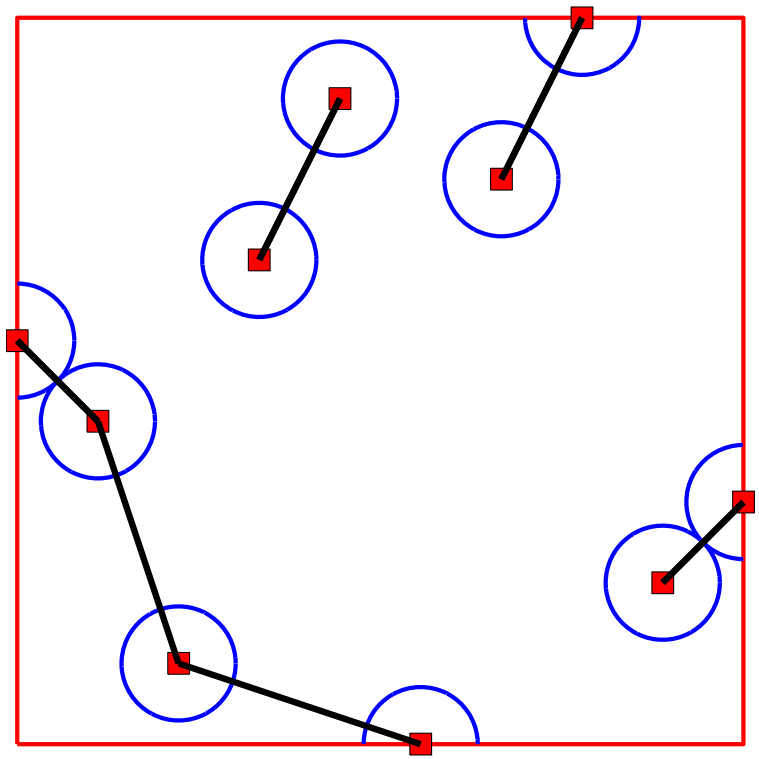


Many results... (Redmond and Yukich, 1996; Yukich, 1998; Penrose and Yukich, 2003; Wade, 2007; Penrose and Yukich, 2011)...

$$\frac{\sum_{e_i \in \mathcal{G}(\mathbf{X}_n)} |e_i|^\beta}{n^{1-\beta/d}} \to C(\beta, d) \int \varphi^{1-\beta/d}(\mathbf{x})\, \mathrm{d}\mathbf{x}\,, \quad n \to \infty$$

with $\mathcal{G}(\mathbf{X}_n)$ Minimum Spanning Tree (MST), NN, TS, Voronoï, Delaunay, Sphere of Influence, Gabriel... (different types of convergence (L_p), different conditions on φ and β...)

$\mathbf{X}_n$ a Lh with $n = 10$, $d = 2$: $\mathcal{G}_{NN}(\mathbf{X}_n)$



To summarize: we construct such a graph $\mathcal{G}$ on $\mathbf{X}_n$, then

$$\Phi_{\mathcal{G},\beta}(\mathbf{X}_n) = \frac{\sum_{e_i \in \mathcal{G}(\mathbf{X}_n)} |e_i|^\beta}{n^{1-\beta/d}} \quad \to \quad C(\beta, d) \int \varphi^{1-\beta/d}(\mathbf{x})\, \mathrm{d}\mathbf{x}\,, \;\; n \to \infty$$

$$= C(\beta, d) \int \varphi^{\alpha}(\mathbf{x})\, \mathrm{d}\mathbf{x}\,,$$

To summarize: we construct such a graph $\mathcal{G}$ on $\mathbf{X}_n$, then

$$\Phi_{\mathcal{G},\beta}(\mathbf{X}_n) = \frac{\sum_{e_i \in \mathcal{G}(\mathbf{X}_n)} |e_i|^\beta}{n^{1-\beta/d}} \quad \to \quad C(\beta,d) \int \varphi^{1-\beta/d}(\mathbf{x})\,\mathrm{d}\mathbf{x}\,, \ n \to \infty$$
$$= C(\beta,d) \int \varphi^{\alpha}(\mathbf{x})\,\mathrm{d}\mathbf{x}\,,$$

- yields an estimate of $H_\alpha(\varphi) = \frac{1}{\alpha-1}\left[1 - \int \varphi^\alpha(\mathbf{x})\,\mathrm{d}\mathbf{x}\right]$ for $\alpha = 1 - \beta/d$ (with a condition on β — typically, $\beta > -d$)

To summarize: we construct such a graph $\mathcal{G}$ on $\mathbf{X}_n$, then

$$\Phi_{\mathcal{G},\beta}(\mathbf{X}_n) = \frac{\sum_{e_i \in \mathcal{G}(\mathbf{X}_n)} |e_i|^\beta}{n^{1-\beta/d}} \quad \to \quad C(\beta,d) \int \varphi^{1-\beta/d}(\mathbf{x})\,\mathrm{d}\mathbf{x}\,, \ n \to \infty$$
$$= C(\beta,d) \int \varphi^{\alpha}(\mathbf{x})\,\mathrm{d}\mathbf{x}\,,$$

- yields an estimate of $H_\alpha(\varphi) = \frac{1}{\alpha-1}\left[1 - \int \varphi^\alpha(\mathbf{x})\,\mathrm{d}\mathbf{x}\right]$ for $\alpha = 1-\beta/d$ (with a condition on β — typically, $\beta > -d$)
- Choice of $\mathbf{X}_n$? maximize H_α with $\alpha > 0$

  if $\alpha > 1$ $(-d < \beta < 0)$ ⇛ minimize $\int \varphi^\alpha(\mathbf{x})\,\mathrm{d}\mathbf{x}$

  ⇛ minimize $\Phi_{\mathcal{G},\beta}(\mathbf{X}_n)$

  For $\mathcal{G}_{NN}$ ⇛ maximize $\overline{\Phi}_{[NN,q]}(\mathbf{X}_n)$ with $0 < q = -\beta < d$

= maximize an intra-distances criterion

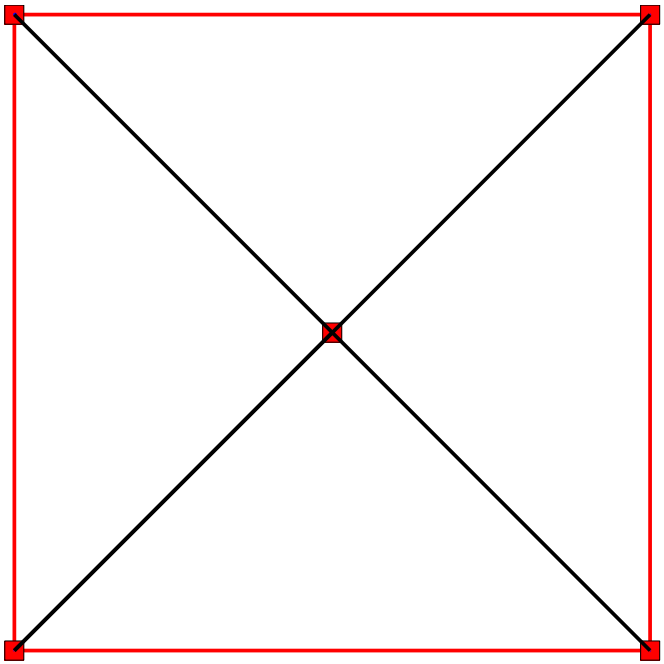
▲ Maximizing $\overline{\Phi}_{[NN,q]}(\mathbf{X}_n)$ with $q < 0$ is not always convenient ▲
(= maximize $\Phi_{\mathcal{G}_{NN},\beta}(\mathbf{X}_n)$ with $\beta > 0$)

▲ Maximizing $\overline{\Phi}_{[NN,q]}(\mathbf{X}_n)$ with $q < 0$ is not always convenient ▲
(= maximize $\Phi_{\mathcal{G}_{NN},\beta}(\mathbf{X}_n)$ with $\beta > 0$)

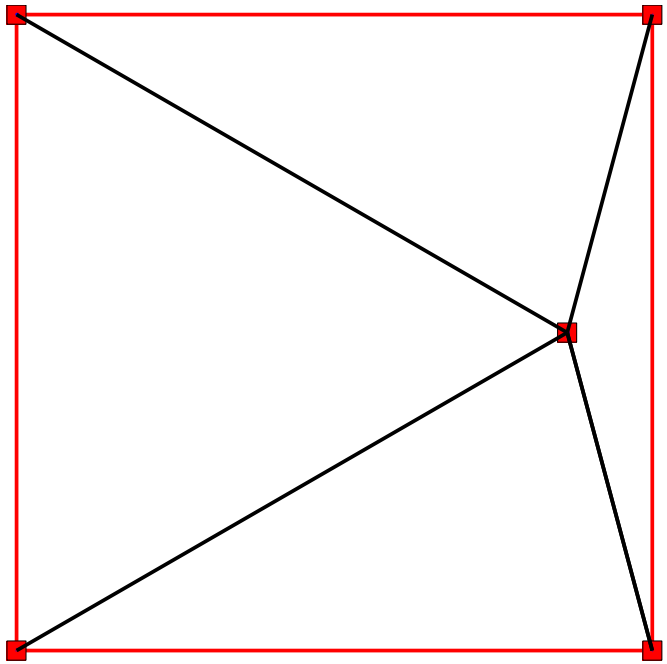
Ex: $\mathscr{X} = [0,1]^2$, comparaison between 2 designs $\mathbf{X}_n^a$ and $\mathbf{X}_n^b$ for $\overline{\Phi}_{[NN,q]}$ with $q = -1$ ($\beta = 1$)

$\mathbf{X}_n^a$

$(1/n)\sum_{e_i \in \mathcal{G}_{NN}(\mathbf{X}_n^a)} |e_i| = \frac{\sqrt{2}}{2} \simeq 0.70711$



$\mathbf{X}_n^b$
central point at $(\sqrt{3}/2, 1/2)$
$(1/n)\sum_{e_i \in \mathcal{G}_{NN}(\mathbf{X}_n^b)} |e_i| \simeq 0.71058$



. . . but the $|e_i|$ have a larger variance for $\mathbf{X}_n^b$ than for $\mathbf{X}_n^a$

MST Graph: Franco (2008); Franco et al. (2009) use a representation in the plane defined by

the mean $E_n = (1/n) \sum_{e_i \in \mathcal{G}_{MST}(\mathbf{X})} |e_i|$ and

the standard deviation $S_n = \left(\text{var}_{\mathcal{G}_{MST}(\mathbf{X})} \{|e_i|\} \right)^{1/2}$

to classify different sorts of space-filling designs $\mathbf{X}_n$

$\rightarrow$ we wish to have a large E_n and a small S_n

MST Graph: Franco (2008); Franco et al. (2009) use a representation in the plane defined by

the mean $E_n = (1/n) \sum_{e_i \in \mathcal{G}_{MST}(\mathbf{X})} |e_i|$ and

the standard deviation $S_n = \left(\text{var}_{\mathcal{G}_{MST}(\mathbf{X})} \{|e_i|\} \right)^{1/2}$

to classify different sorts of space-filling designs $\mathbf{X}_n$

→ we wish to have a large E_n and a small S_n

We might use another graph $\mathcal{G}$ than $\mathcal{G}_{MST}$ (e.g., $\mathcal{G}_{NN}$)

and use $|e_i|^\beta$ (e.g., with $-d < \beta < 0$)

MST Graph: Franco (2008); Franco et al. (2009) use a representation in the plane defined by

the mean $E_n = (1/n) \sum_{e_i \in \mathcal{G}_{MST}(\mathbf{X})} |e_i|$ and

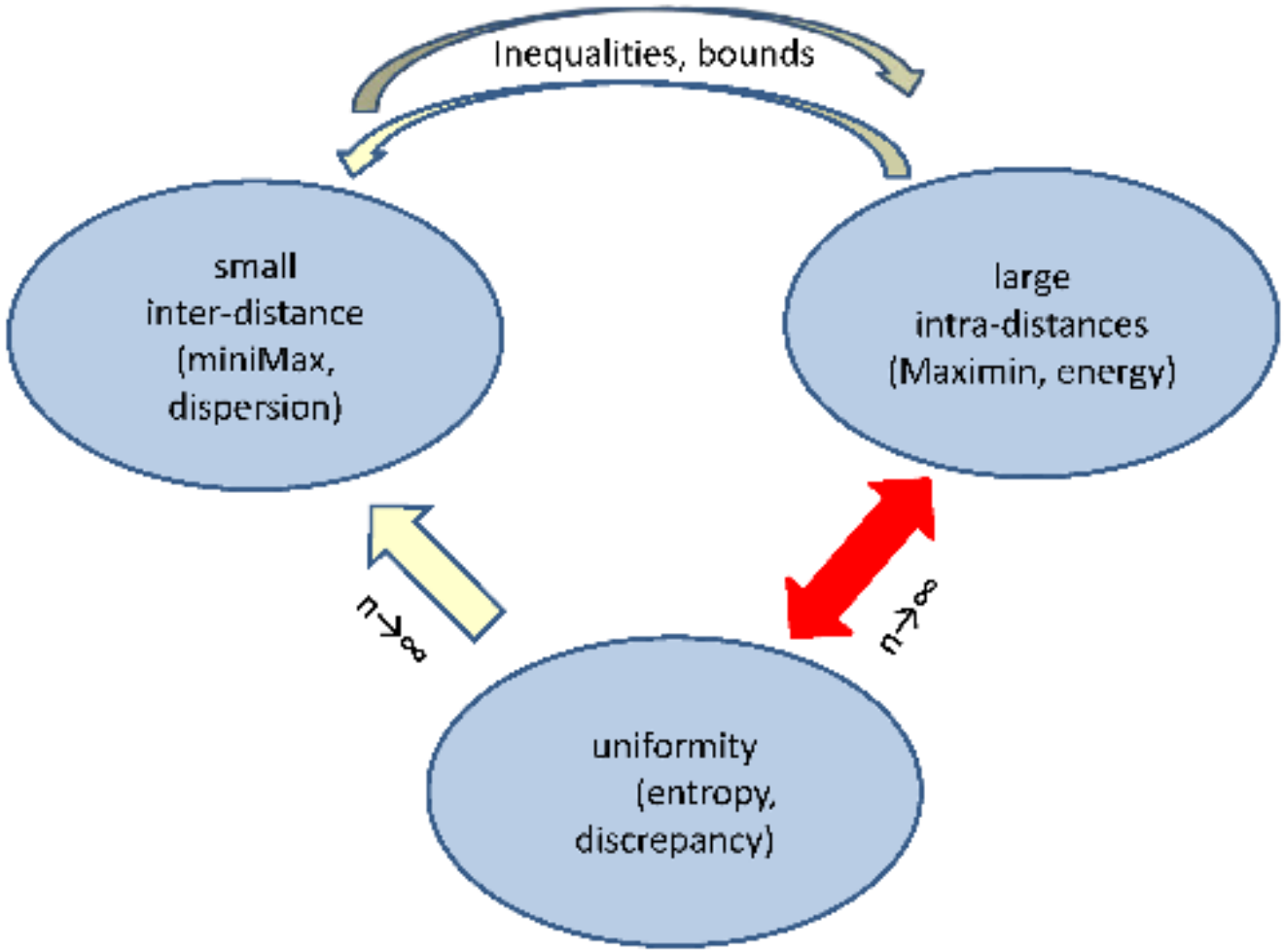
the standard deviation $S_n = \left(\text{var}_{\mathcal{G}_{MST}(\mathbf{X})} \{|e_i|\} \right)^{1/2}$

to classify different sorts of space-filling designs $\mathbf{X}_n$

➞ we wish to have a large E_n and a small S_n

We might use another graph $\mathcal{G}$ than $\mathcal{G}_{MST}$ (e.g., $\mathcal{G}_{NN}$)
and use $|e_i|^\beta$ (e.g., with $-d < \beta < 0$)

➠ For $\Phi_{Mm}(\cdot)$, or $\overline{\Phi}_{[q]}(\cdot)$ with $q > d$, or $\overline{\Phi}_{[NN,q]}(\cdot)$ with $q > 0$,
the distribution of an optimum design should be close to uniformity


Inequalities, bounds
small
inter-distance
(miniMax,
dispersion)
large
intra-distances
(Maximin, energy)
uniformity
(entropy,
discrepancy)
$n \to \infty$
$n \to \infty$

2.2 **Discrepancy: motivation** (see (Niederreiter, 1992, Chap. 1,2) — true monument with 371 references)

Integration with Monte Carlo (MC) method

$\mathscr{X}$ compact $\subset \mathbb{R}^d$,

$$\int_{\mathscr{X}} f(\mathbf{u})\,\mathrm{d}\mathbf{u} \quad \simeq \quad \mathrm{vol}(\mathscr{X})\frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

for x_i i.i.d. $\sim \mu$ uniform over $\mathscr{X}$ ➠ error $\simeq \mathcal{O}(n^{-1/2})$

2.2 **Discrepancy: motivation** (see (Niederreiter, 1992, Chap. 1,2) — true monument with 371 references)

Integration with Monte Carlo (MC) method

$\mathscr{X}$ compact $\subset \mathbb{R}^d$,

$$\int_{\mathscr{X}} f(\mathbf{u})\,\mathrm{d}\mathbf{u} \quad \simeq \quad \mathrm{vol}(\mathscr{X})\frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

for x_i i.i.d. $\sim \mu$ uniform over $\mathscr{X}$ ➠ error $\simeq \mathcal{O}(n^{-1/2})$

Trapezoidal rule in dimension d ➠ error $\simeq \mathcal{O}(n^{-2/d})$

➠ MC better than trapezoidal rule for $d \geq 5$ (without any regularity assumption on f)

We can do better: quasi-Monte Carlo (QMC) method

➠ discrepancy

Quasi-Monte Carlo (QMC) method

Evaluate f at deterministic $\mathbf{x}_i$ in $\mathscr{X} = \mathbb{I}_d \triangleq [0,1]^d$:

$$\widehat{I}_n \triangleq \frac{1}{n}\sum_{i=1}^n f(\mathbf{x}_i) \to I(f) = \int_{\mathbb{I}_d} f(\mathbf{u})\,\mathrm{d}(\mathbf{u}),\ n \to \infty,$$

for all Riemann integrable f if the $\mathbf{x}_1, \mathbf{x}_2, \ldots$ are uniformly distributed in $\mathbb{I}_d$

Quasi-Monte Carlo (QMC) method

Evaluate f at deterministic $\mathbf{x}_i$ in $\mathscr{X} = \mathbb{I}_d \triangleq [0,1]^d$:

$$\widehat{I}_n \triangleq \frac{1}{n}\sum_{i=1}^n f(\mathbf{x}_i) \to I(f) = \int_{\mathbb{I}_d} f(\mathbf{u})\,\mathrm{d}(\mathbf{u}),\ n \to \infty,$$

for all Riemann integrable f if the $\mathbf{x}_1, \mathbf{x}_2, \ldots$ are uniformly distributed in $\mathbb{I}_d$

➠ requires $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathcal{I}_{\mathbb{B}}(\mathbf{x}) = \mathrm{vol}(\mathbb{B})$ for all $\mathbb{B} \subset \mathbb{I}_d$

Speed of convergence of $\widehat{I}_n \to I(f)$?

The distribution of $\mathbf{x}_i$ must be close to uniform:

discrepancy measures distance to uniformity

Quasi-Monte Carlo (QMC) method

Evaluate f at deterministic $\mathbf{x}_i$ in $\mathscr{X} = \mathbb{I}_d \triangleq [0,1]^d$:

$$\widehat{I}_n \triangleq \frac{1}{n}\sum_{i=1}^n f(\mathbf{x}_i) \to I(f) = \int_{\mathbb{I}_d} f(\mathbf{u})\,\mathrm{d}(\mathbf{u}),\ n \to \infty,$$

for all Riemann integrable f if the $\mathbf{x}_1, \mathbf{x}_2, \ldots$ are uniformly distributed in $\mathbb{I}_d$

➠ requires $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathcal{I}_{\mathbb{B}}(\mathbf{x}) = \mathrm{vol}(\mathbb{B})$ for all $\mathbb{B} \subset \mathbb{I}_d$

Speed of convergence of $\widehat{I}_n \to I(f)$?

The distribution of $\mathbf{x}_i$ must be close to uniform:

discrepancy measures distance to uniformity

Discrepancy

$$\triangleright \quad D_n(\mathscr{B}, \mathbf{X}_n) \triangleq \sup_{\mathbb{B}\in\mathscr{B}} \left| \frac{\text{nb. of } \mathbf{x}_i \text{ in } \mathbb{B}}{n} - \mathrm{vol}(\mathbb{B}) \right| \quad \triangleleft$$

with $\mathscr{B}$ a family of subsets of $\mathbb{I}_d$ ($\Rightarrow 0 \leq D_n(\mathscr{B}, \mathbf{X}_n) \leq 1$)

Quasi-Monte Carlo (QMC) method

Evaluate f at deterministic $\mathbf{x}_i$ in $\mathscr{X} = \mathbb{I}_d \triangleq [0,1]^d$:

$$\widehat{I}_n \triangleq \frac{1}{n}\sum_{i=1}^n f(\mathbf{x}_i) \to I(f) = \int_{\mathbb{I}_d} f(\mathbf{u})\,\mathrm{d}(\mathbf{u}),\ n \to \infty,$$

for all Riemann integrable f if the $\mathbf{x}_1, \mathbf{x}_2, \ldots$ are uniformly distributed in $\mathbb{I}_d$

➠ requires $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathcal{I}_{\mathbb{B}}(\mathbf{x}) = \mathrm{vol}(\mathbb{B})$ for all $\mathbb{B} \subset \mathbb{I}_d$

Speed of convergence of $\widehat{I}_n \to I(f)$?

The distribution of $\mathbf{x}_i$ must be close to uniform:

discrepancy measures distance to uniformity

Discrepancy

$$\triangleright \quad D_n(\mathscr{B}, \mathbf{X}_n) \triangleq \sup_{\mathbb{B}\in\mathscr{B}} \left| \frac{\text{nb. of } \mathbf{x}_i \text{ in } \mathbb{B}}{n} - \mathrm{vol}(\mathbb{B}) \right| \quad \triangleleft$$

with $\mathscr{B}$ a family of subsets of $\mathbb{I}_d$ ($\Rightarrow 0 \leq D_n(\mathscr{B}, \mathbf{X}_n) \leq 1$)

We shall consider particular families $\mathscr{B}$ ...

Two important special cases:

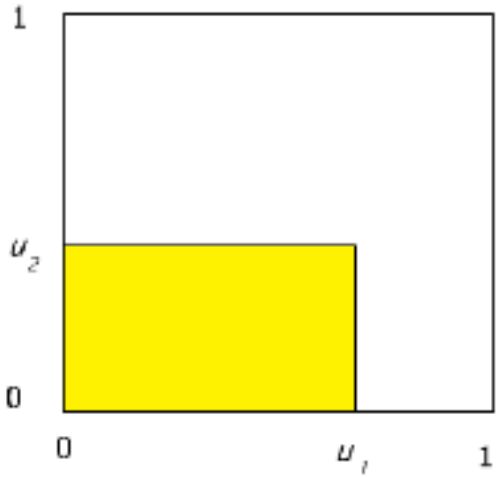
- Star-discrepancy $D_n^*(\mathbf{X}_n) = D_n(\mathscr{B}, \mathbf{X}_n)$ when
  $\mathscr{B}$ contains all subsets defined by $\prod_{\ell=1}^d [0, u_\ell)$
  $\boxed{D_n^*(\mathbf{X}_n) = \sup_{\mathbf{u}\in[0,1]^d} |\mathrm{F}_n(\mathbf{u}) - \mathrm{F}_U(\mathbf{u})|}$
  ($\mathrm{F}_n(\cdot) \triangleq$ empirical d.d.f., $\mathrm{F}_U(\mathbf{u}) \triangleq \prod_{\ell=1}^d \{\mathbf{u}\}_\ell =$ c.d.f. of uniform)
- Extreme discrepancy $D_n(\mathbf{X}_n) = D_n(\mathscr{B}, \mathbf{X}_n)$ when
  $\mathscr{B}$ contains all subsets defined by $\prod_{\ell=1}^d [u_\ell, v_\ell)$

For any $\mathbf{X}_n \subset \mathbb{I}_d^n$, we have $D_n^*(\mathbf{X}_n) \leq D_n(\mathbf{X}_n) \leq 2^d D_n^*(\mathbf{X}_n)$

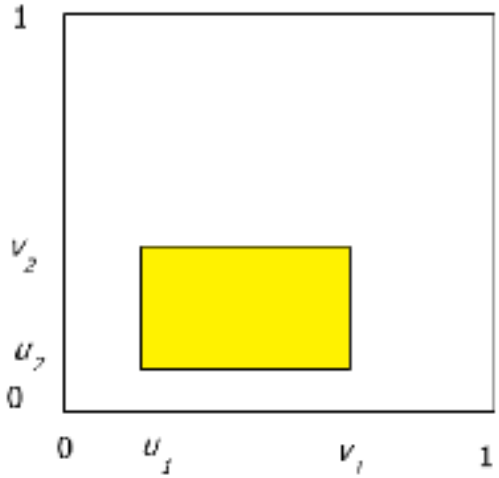
Two important special cases:

- Star-discrepancy $D_n^*(\mathbf{X}_n) = D_n(\mathscr{B}, \mathbf{X}_n)$ when
  $\mathscr{B}$ contains all subsets defined by $\prod_{\ell=1}^d [0, u_\ell)$
  $$\boxed{D_n^*(\mathbf{X}_n) = \sup_{\mathbf{u} \in [0,1]^d} |F_n(\mathbf{u}) - F_U(\mathbf{u})|}$$
  ($F_n(\cdot) \triangleq$ empirical d.d.f., $F_U(\mathbf{u}) \triangleq \prod_{\ell=1}^d \{\mathbf{u}\}_\ell =$ c.d.f. of uniform)
- Extreme discrepancy $D_n(\mathbf{X}_n) = D_n(\mathscr{B}, \mathbf{X}_n)$ when
  $\mathscr{B}$ contains all subsets defined by $\prod_{\ell=1}^d [u_\ell, v_\ell)$

For any $\mathbf{X}_n \subset \mathbb{I}_d^n$, we have $D_n^*(\mathbf{X}_n) \leq D_n(\mathbf{X}_n) \leq 2^d D_n^*(\mathbf{X}_n)$



$D_n^*(\mathbf{X}_n)$: $\mathbb{B} = \prod_{\ell=1}^d [0, u_\ell)$



$D_n(\mathbf{X}_n)$: $\mathbb{B} = \prod_{\ell=1}^d [u_\ell, v_\ell)$

In dimension $d = 1$, with $0 \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq 1$:

$$D_n^*(\mathbf{X}_n) = \frac{1}{2n} + \max_{1 \leq i \leq n} \left| x_i - \frac{2i-1}{2n} \right|$$

(➠ Kolmogorov-Smirnov test for uniformity)

$$D_n(\mathbf{X}_n) = \frac{1}{n} + \max_{1 \leq i \leq n} \left(\frac{i}{n} - x_i \right) - \min_{1 \leq i \leq n} \left(\frac{i}{n} - x_i \right)$$

and $D_n^*(\mathbf{X}_n) \geq \frac{1}{2n}$, $D_n(\mathbf{X}_n) \geq \frac{1}{n}$, with equality for $\mathbf{X}_{mM,n}^* : x_i = \frac{2i-1}{2n} \ \forall i$

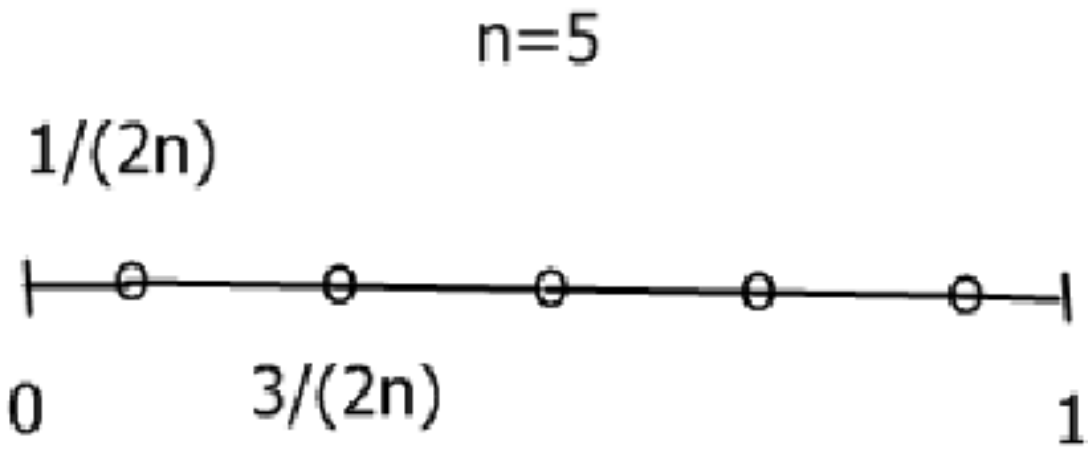
In dimension $d = 1$, with $0 \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq 1$:

$$D_n^*(\mathbf{X}_n) = \frac{1}{2n} + \max_{1 \leq i \leq n} \left| x_i - \frac{2i-1}{2n} \right|$$

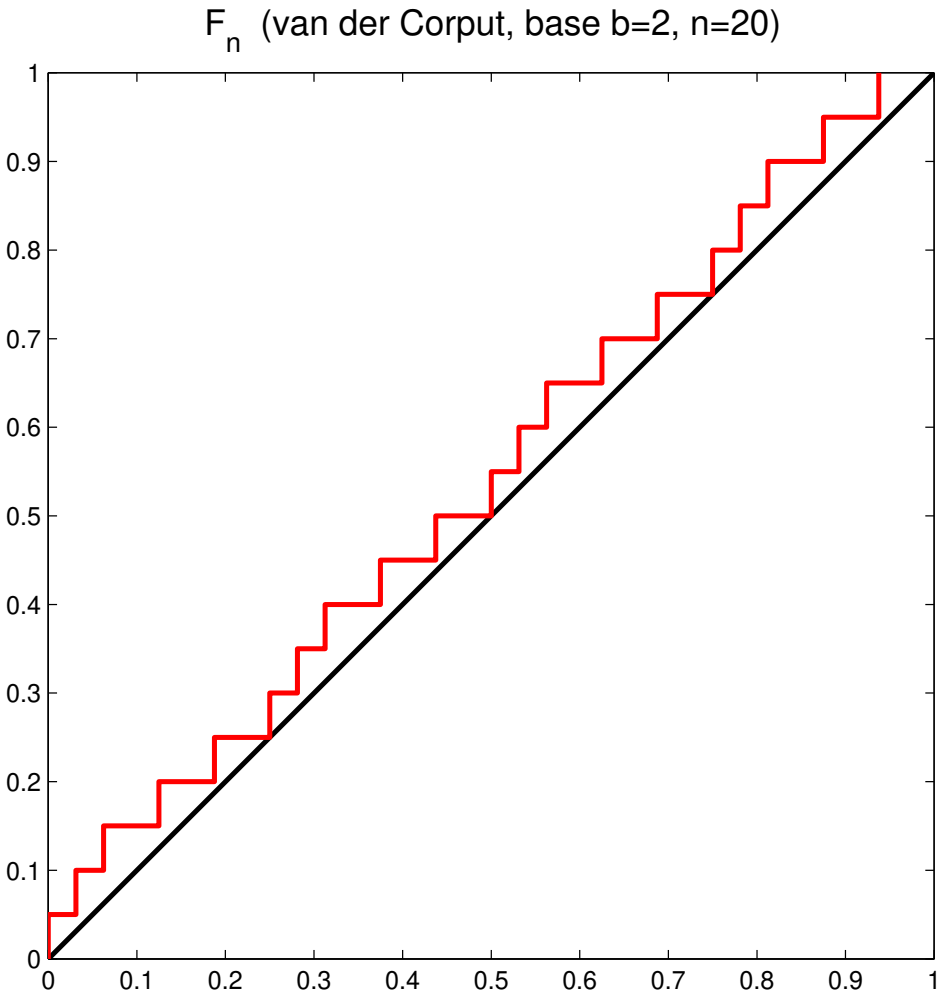
(➠ Kolmogorov-Smirnov test for uniformity)

$$D_n(\mathbf{X}_n) = \frac{1}{n} + \max_{1 \leq i \leq n} \left(\frac{i}{n} - x_i \right) - \min_{1 \leq i \leq n} \left(\frac{i}{n} - x_i \right)$$

and $D_n^*(\mathbf{X}_n) \geq \frac{1}{2n}$, $D_n(\mathbf{X}_n) \geq \frac{1}{n}$, with equality for $\mathbf{X}_{mM,n}^* : x_i = \frac{2i-1}{2n}\ \forall i$



empirical c.d.f. $F_n(x)$ and $F_U(x)$
$d = 1$, $\mathbf{X}_{20}$ = first 20 points of van der Corput sequence in base 2
(voir § 2.4)



Why is it important?

For $d = 1$, f with bounded variation on $[0,1]$ ($V(f) \triangleq \int_0^1 |\mathrm{d}f(u)| < \infty$)

$$\boxed{\left| \int_0^1 f(u)\,\mathrm{d}u - \frac{1}{n}\sum_{i=1}^n f(x_i) \right| \leq D_n^*(\mathbf{X}_n)\, V(f)}$$

➠ Koksma (1942/1943) inequality (cannot be improved)
(easy proof, integration by parts)

Therefore, for $d = 1$

$$\left| \int_0^1 f(u)\,\mathrm{d}u - \frac{1}{n}\sum_{i=1}^n f(x_i) \right| \leq \frac{V(f)}{2n} \quad \text{for } \mathbf{X}_n^* \text{ such that } x_i = \frac{2i-1}{2n}\ \forall i$$

whereas MC error $\simeq \frac{\sigma(f)}{n^{1/2}}$

Therefore, for $d = 1$

$$\left| \int_0^1 f(u)\,\mathrm{d}u - \frac{1}{n}\sum_{i=1}^n f(x_i) \right| \leq \frac{V(f)}{2n}$$ for $\mathbf{X}_n^*$ such that $x_i = \frac{2i-1}{2n}\ \forall i$

whereas MC error $\simeq \frac{\sigma(f)}{n^{1/2}}$

In dimension $d \geq 2$: $$\left| \int_{\mathbb{I}_d} f(\mathbf{u})\,\mathrm{d}\mathbf{u} - \frac{1}{n}\sum_{i=1}^n f(\mathbf{x}_i) \right| \leq D_n^*(\mathbf{X}_n)\, V(f)$$

(Koksma-Hlawka (1961) inequality, cannot be improved)
with $V(f)$ = variation in the sense of Hardy and Krause, and

$$\left| \int_{\mathbb{I}_d} f(\mathbf{u})\,\mathrm{d}\mathbf{u} - \frac{1}{n}\sum_{i=1}^n f(\mathbf{x}_i) \right| \leq \begin{cases} V(f)\, C_d\, \dfrac{(\log n)^{d-1}}{n} \\ \text{for } \mathbf{X}_n = \text{Hammersley point set} \\ V(f)\, C_d'\, \dfrac{(\log n)^{d}}{n} \\ \text{for } \mathbf{X}_n = \text{first } n \text{ elements} \\ \text{of Halton sequence (for instance) } \mathbf{X}_\infty \end{cases}$$

voir § 2.4

Therefore, for $d = 1$

$$\left| \int_0^1 f(u)\,\mathrm{d}u - \frac{1}{n}\sum_{i=1}^n f(x_i) \right| \leq \frac{V(f)}{2n}$$ for $\mathbf{X}_n^*$ such that $x_i = \frac{2i-1}{2n}\ \forall i$

whereas MC error $\simeq \frac{\sigma(f)}{n^{1/2}}$

In dimension $d \geq 2$: $$\left| \int_{\mathbb{I}_d} f(\mathbf{u})\,\mathrm{d}\mathbf{u} - \frac{1}{n}\sum_{i=1}^n f(\mathbf{x}_i) \right| \leq D_n^*(\mathbf{X}_n)\, V(f)$$

(Koksma-Hlawka (1961) inequality, cannot be improved)
with $V(f)$ = variation in the sense of Hardy and Krause, and

$$\left| \int_{\mathbb{I}_d} f(\mathbf{u})\,\mathrm{d}\mathbf{u} - \frac{1}{n}\sum_{i=1}^n f(\mathbf{x}_i) \right| \leq \begin{cases} V(f)\, C_d\, \dfrac{(\log n)^{d-1}}{n} \\ \text{for } \mathbf{X}_n = \text{Hammersley point set} \\ V(f)\, C_d'\, \dfrac{(\log n)^{d}}{n} \\ \text{for } \mathbf{X}_n = \text{first } n \text{ elements} \\ \text{of Halton sequence (for instance) } \mathbf{X}_\infty \end{cases}$$

voir § 2.4

- Error $\searrow$ faster than for MC, but which constant C_d ?
- n-point set $\mathbf{X}_n \neq$ first n elements of an infinite sequence $\mathbf{X}_\infty$

2.3 Discrepancy criteria

Difficulty : $D_n^*(\mathbf{X}_n)$ and $D_n(\mathbf{X}_n)$ are difficult to compute for $d \geq 2$

see, e.g., (Dobkin and Eppstein, 1993; Thiémard, 2001; Gnewuch et al., 2012) and the references therein

➠ (many!) other definitions of discrepancy

2.3 Discrepancy criteria

Difficulty : $D_n^*(\mathbf{X}_n)$ and $D_n(\mathbf{X}_n)$ are difficult to compute for $d \geq 2$
see, e.g., (Dobkin and Eppstein, 1993; Thiémard, 2001; Gnewuch et al., 2012) and the references therein

➠ (many!) other definitions of discrepancy

One may wish to have:

① invariance by permutation of principal axes, by reflection w.r.t. center of $\mathbb{I}_d$

② a uniformity property on d' dimensional subspaces, $d' < d$

③ a geometrical interpretation

④ a sort of Koksma-Hlawka inequality

⑤ . . . and easy evaluation!

For ⑤: substitute a L_p norm for the L_∞ norm

$$D_n^*(\mathbf{X}_n) = \sup_{\mathbf{u}\in[0,1]^d} |F_n(\mathbf{u}) - F_U(\mathbf{u})| \Rightarrow \left(\int_{[0,1]^d} |F_n(\mathbf{u}) - F_U(\mathbf{u})|^p \, \mathrm{d}\mathbf{u} \right)^{1/p}$$

⇛ Analytical expression for $p = 2$, fine for ③, ④ and ⑤, but not for ① and ②

For ⑤: substitute a L_p norm for the L_∞ norm

$D_n^*(\mathbf{X}_n) = \sup_{\mathbf{u}\in[0,1]^d} |\mathrm{F}_n(\mathbf{u}) - \mathrm{F}_U(\mathbf{u})|$ ➠ $\left(\int_{[0,1]^d} |\mathrm{F}_n(\mathbf{u}) - \mathrm{F}_U(\mathbf{u})|^p \,\mathrm{d}\mathbf{u}\right)^{1/p}$

➠ Analytical expression for $p = 2$, fine for ③, ④ and ⑤, but not for ① and ②

For ②: consider projections on all d' dimensional faces, $d' < d$

$D_n^*(\mathbf{X}_n)$ ➠

$$\left(\sum_{d'=1}^{d} \sum_{i_1<\cdots<i_{d'}} \int_{[0,1]^{d'}} |\mathrm{F}_n(\{\mathbf{u}\}_{i_1,\ldots,i_{d'}}) - \mathrm{F}_U(\{\mathbf{u}\}_{i_1,\ldots,i_{d'}})|^p \,\mathrm{d}\{\mathbf{u}\}_{i_1,\ldots,i_{d'}}\right)^{1/p}$$

➠ Analytical expression for $p = 2$, fine for ②, ③, ④ and ⑤, but not for ①

For ⑤: substitute a L_p norm for the L_∞ norm

$D_n^*(\mathbf{X}_n) = \sup_{\mathbf{u}\in[0,1]^d} |\mathrm{F}_n(\mathbf{u}) - \mathrm{F}_U(\mathbf{u})| \rightarrow \left(\int_{[0,1]^d} |\mathrm{F}_n(\mathbf{u}) - \mathrm{F}_U(\mathbf{u})|^p \,\mathrm{d}\mathbf{u}\right)^{1/p}$

➠ Analytical expression for $p = 2$, fine for ③, ④ and ⑤, but not for ① and ②

For ②: consider projections on all d' dimensional faces, $d' < d$

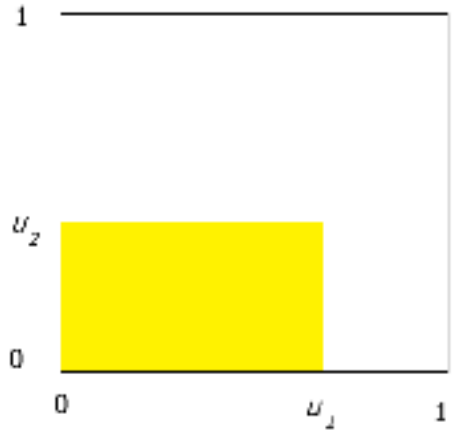
$D_n^*(\mathbf{X}_n)$ ➠

$$\left(\sum_{d'=1}^{d} \sum_{i_1<\cdots<i_{d'}} \int_{[0,1]^{d'}} |\mathrm{F}_n(\{\mathbf{u}\}_{i_1,\ldots,i_{d'}}) - \mathrm{F}_U(\{\mathbf{u}\}_{i_1,\ldots,i_{d'}})|^p \,\mathrm{d}\{\mathbf{u}\}_{i_1,\ldots,i_{d'}}\right)^{1/p}$$

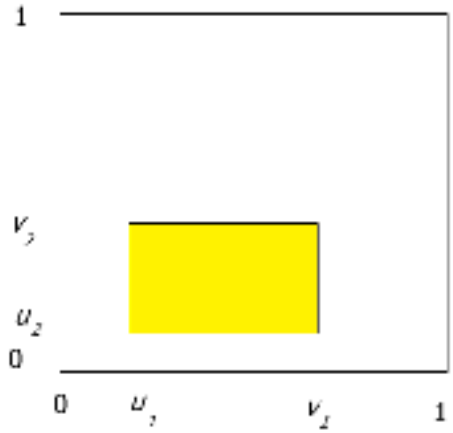
➠ Analytical expression for $p = 2$, fine for ②, ③, ④ and ⑤, but not for ①

For ①: change the family of sets $\mathbb{B}$ in calculation of discrepancy

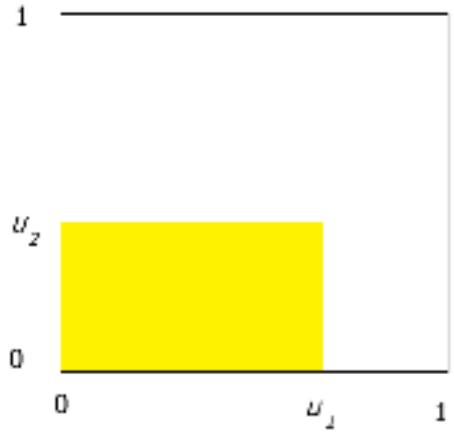
$D_n^*(\mathbf{X}_n)$: $\mathbb{B} = \prod_{\ell=1}^d [0, u_\ell)$



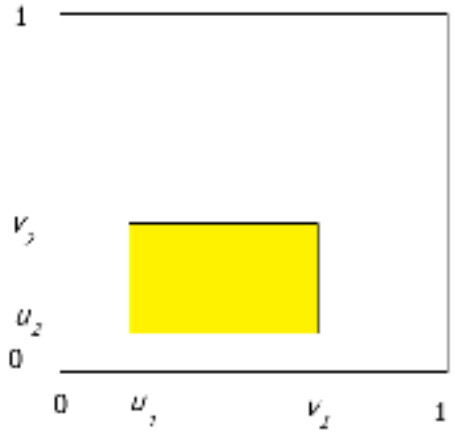
$D_n(\mathbf{X}_n)$: $\mathbb{B} = \prod_{\ell=1}^d [u_\ell, v_\ell)$



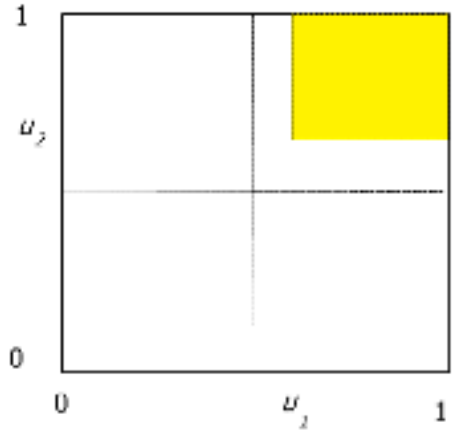
$D_n^*(\mathbf{X}_n)$: $\mathbb{B} = \prod_{\ell=1}^d [0, u_\ell)$



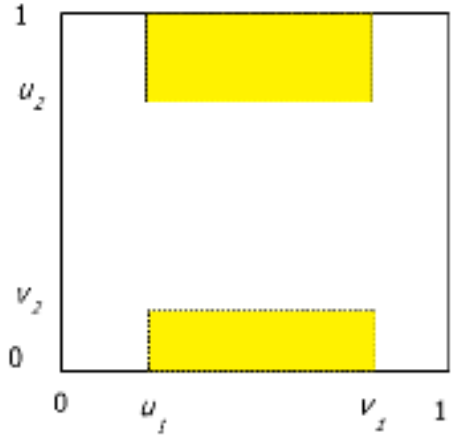
$D_n(\mathbf{X}_n)$: $\mathbb{B} = \prod_{\ell=1}^d [u_\ell, v_\ell)$



Centered discrepancy: consider the vertex of the (d' dimensional) cube closest to $\mathbf{u}$



Wrap-around discrepancy: consider $[u_i, v_i]$ if $u_i \leq v_i$ and $[u_i, 1] \cup [0, v_i]$ otherwise



They can be computed...

$$D_{Cent,L_2}(\mathbf{X}_n) = \left[\left(\frac{13}{12}\right)^d - \frac{2}{n}\sum_{k=1}^{n}\prod_{i=1}^{d}\left(1+\frac{1}{2}\left|\{\mathbf{x}_k\}_i-\frac{1}{2}\right|-\frac{1}{2}\left|\{\mathbf{x}_k\}_i-\frac{1}{2}\right|^2\right)\right.$$
$$\left.+\frac{1}{n^2}\sum_{k,k'=1}^{n}\prod_{i=1}^{d}\left(1+\frac{1}{2}\left|\{\mathbf{x}_k\}_i-\frac{1}{2}\right|+\frac{1}{2}\left|\{\mathbf{x}_{k'}\}_i-\frac{1}{2}\right|-\frac{1}{2}\left|\{\mathbf{x}_k\}_i-\{\mathbf{x}_{k'}\}_i\right|\right)\right]^{1/2}$$

$$D_{WA,L_2}(\mathbf{X}_n) = \left\{\frac{1}{n^2}\sum_{k,k'=1}^{n}\prod_{i=1}^{d}\left[\frac{3}{2}-\left|\{\mathbf{x}_k\}_i-\{\mathbf{x}_{k'}\}_i\right|\left(1-\left|\{\mathbf{x}_k\}_i-\{\mathbf{x}_{k'}\}_i\right|\right)\right]\right.$$
$$\left.-\left(\frac{4}{3}\right)^d\right\}^{1/2}$$

see Hickernell (1998a,b); Fang and Ma (2001)
they are differentiable w.r.t. $\mathbf{X}_n$ and can be minimized (Fang and Ma, 2001; Fang et al., 2003, 2005)

They can be computed...

$$D_{Cent,L_2}(\mathbf{X}_n) = \left[\left(\frac{13}{12}\right)^d - \frac{2}{n}\sum_{k=1}^{n}\prod_{i=1}^{d}\left(1+\frac{1}{2}\left|\{\mathbf{x}_k\}_i-\frac{1}{2}\right|-\frac{1}{2}\left|\{\mathbf{x}_k\}_i-\frac{1}{2}\right|^2\right)\right.$$

$$\left.+\frac{1}{n^2}\sum_{k,k'=1}^{n}\prod_{i=1}^{d}\left(1+\frac{1}{2}\left|\{\mathbf{x}_k\}_i-\frac{1}{2}\right|+\frac{1}{2}\left|\{\mathbf{x}_{k'}\}_i-\frac{1}{2}\right|-\frac{1}{2}\left|\{\mathbf{x}_k\}_i-\{\mathbf{x}_{k'}\}_i\right|\right)\right]^{1/2}$$

$$D_{WA,L_2}(\mathbf{X}_n) = \left\{\frac{1}{n^2}\sum_{k,k'=1}^{n}\prod_{i=1}^{d}\left[\frac{3}{2}-\left|\{\mathbf{x}_k\}_i-\{\mathbf{x}_{k'}\}_i\right|\left(1-\left|\{\mathbf{x}_k\}_i-\{\mathbf{x}_{k'}\}_i\right|\right)\right]\right.$$

$$\left.-\left(\frac{4}{3}\right)^d\right\}^{1/2}$$

see Hickernell (1998a,b); Fang and Ma (2001)

they are differentiable w.r.t. $\mathbf{X}_n$ and can be minimized (Fang and Ma, 2001; Fang et al., 2003, 2005)

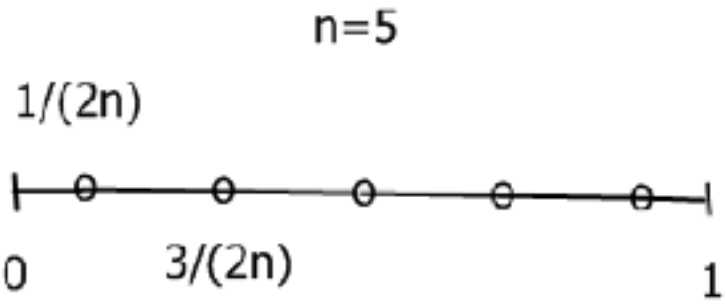
➠ However, generating low discrepancy sequences of points is much easier!

2.4 **Low discrepancy sequences (LDS)** (Niederreiter, 1992, Chap. 3)

$\mathbf{X}_n$ has low discrepancy if $D_n^*(\mathbf{X}_n^*)$ is small
(or $D_n(\mathbf{X}_n^*)$, or another discrepancy)

In dimension 1:

$\mathbf{X}_n^*$ such that $x_i = \frac{2i-1}{2n}$, $i = 1, \ldots, n$



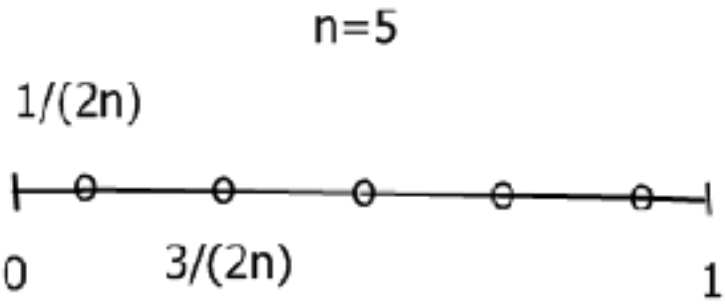
gives $n \times D_n^*(\mathbf{X}_n^*) = 1/2$ for all n

2.4 **Low discrepancy sequences (LDS)** (Niederreiter, 1992, Chap. 3)

$\mathbf{X}_n$ has low discrepancy if $D_n^*(\mathbf{X}_n^*)$ is small
(or $D_n(\mathbf{X}_n^*)$, or another discrepancy)

In dimension 1:

$\mathbf{X}_n^*$ such that $x_i = \frac{2i-1}{2n}$, $i = 1, \ldots, n$



gives $n \times D_n^*(\mathbf{X}_n^*) = 1/2$ for all n

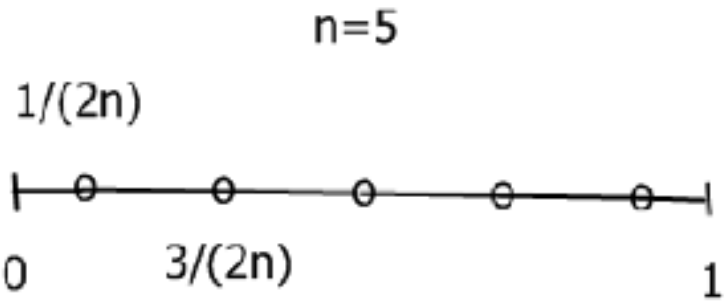
but we cannot construct a sequence $\mathbf{X}_\infty = (x_1, x_2, x_3 \ldots)$ such that n consecutive elements satisfy $n \times D_n^*(\mathbf{X}_\infty) = \mathcal{O}(1)$

2.4 **Low discrepancy sequences (LDS)** (Niederreiter, 1992, Chap. 3)

$\mathbf{X}_n$ has low discrepancy if $D_n^*(\mathbf{X}_n^*)$ is small
(or $D_n(\mathbf{X}_n^*)$, or another discrepancy)

In dimension 1:

$\mathbf{X}_n^*$ such that $x_i = \frac{2i-1}{2n}$, $i = 1, \ldots, n$



gives $n \times D_n^*(\mathbf{X}_n^*) = 1/2$ for all n

but we cannot construct a sequence $\mathbf{X}_\infty = (x_1, x_2, x_3 \ldots)$ such that n consecutive elements satisfy $n \times D_n^*(\mathbf{X}_\infty) = \mathcal{O}(1)$

$n \times D_n^*(\mathbf{X}_\infty)$ necessarily fluctuates: $D_n^*(\mathbf{X}_\infty) > c\frac{\log n}{n}$ infinitely often
(best constant known $c = 0.06$)

We know sequences $\mathbf{X}_\infty$ such that $\boxed{D_n^*(\mathbf{X}_\infty) = \mathcal{O}\left(\frac{\log n}{n}\right)}$

- fractional parts (mainly for $d = 1$)
  for given $n \leftrightarrow$ Lattices
- van der Corput sequences ($d = 1$)
- ➠ Halton sequences ($d > 1$)
- (t, m, d)-nets and (t, d)-sequences (Sobol', Faure)

⏩ → van der Corput

Fractional parts

For $d = 1$: $\boxed{x_k \triangleq \{kz\} = kz - \lfloor kz \rfloor}$, $k = 1, 2 \ldots$, with z irrational (fractional part of kz)

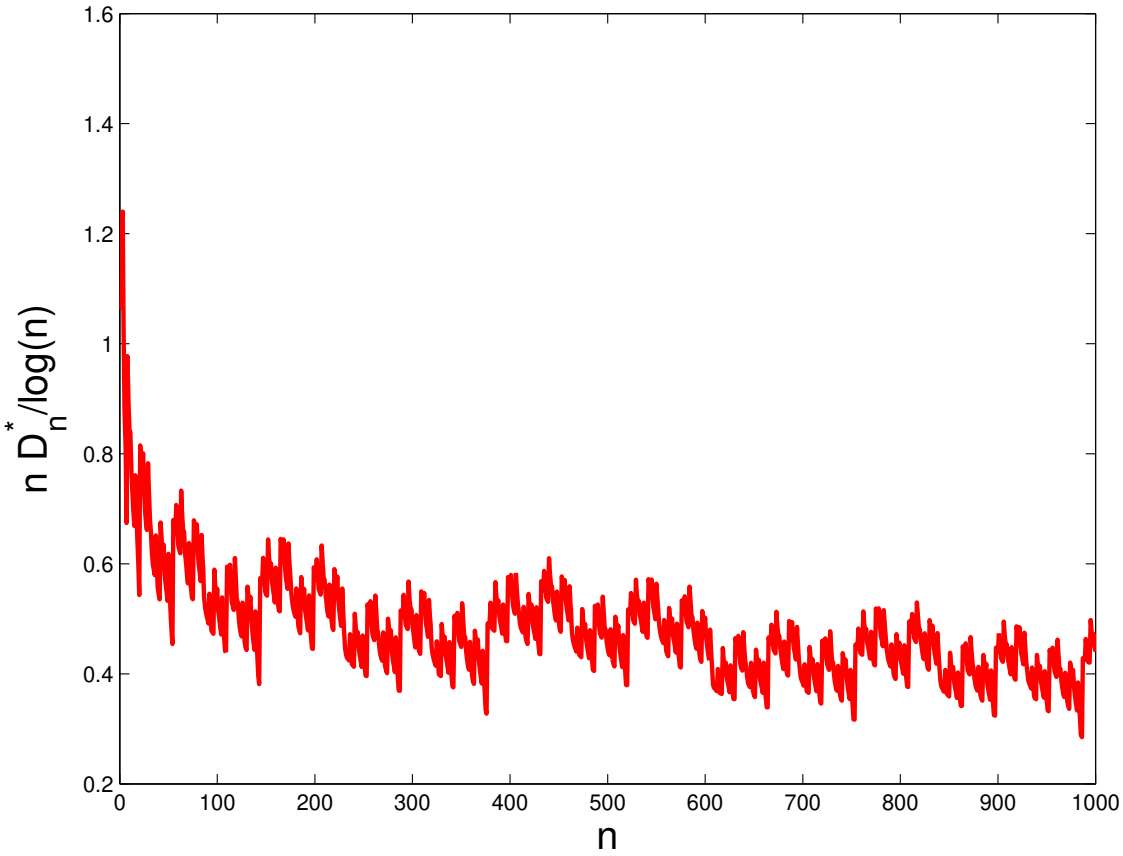
For instance, $z = \varphi = (\sqrt{5} + 1)/2 \simeq 1.618034 =$ Golden section

Fractional parts

For $d = 1$: $\boxed{x_k \triangleq \{kz\} = kz - \lfloor kz \rfloor}$, $k = 1, 2 \ldots$, with z irrational (fractional part of kz)

For instance, $z = \varphi = (\sqrt{5}+1)/2 \simeq 1.618034 =$ Golden section

$(n/\log n) \times D_n^*$

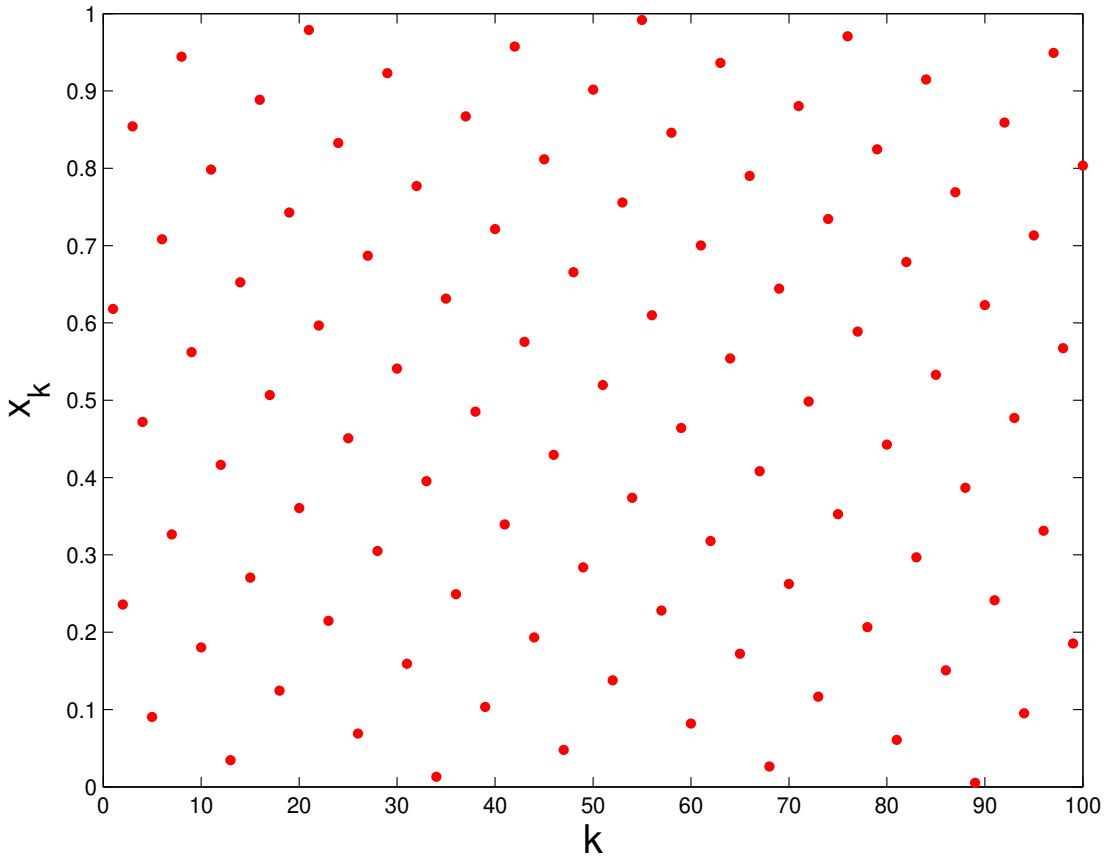


Fractional parts

For $d = 1$: $\boxed{x_k \triangleq \{kz\} = kz - \lfloor kz \rfloor}$, $k = 1, 2 \ldots$, with z irrational (fractional part of kz)

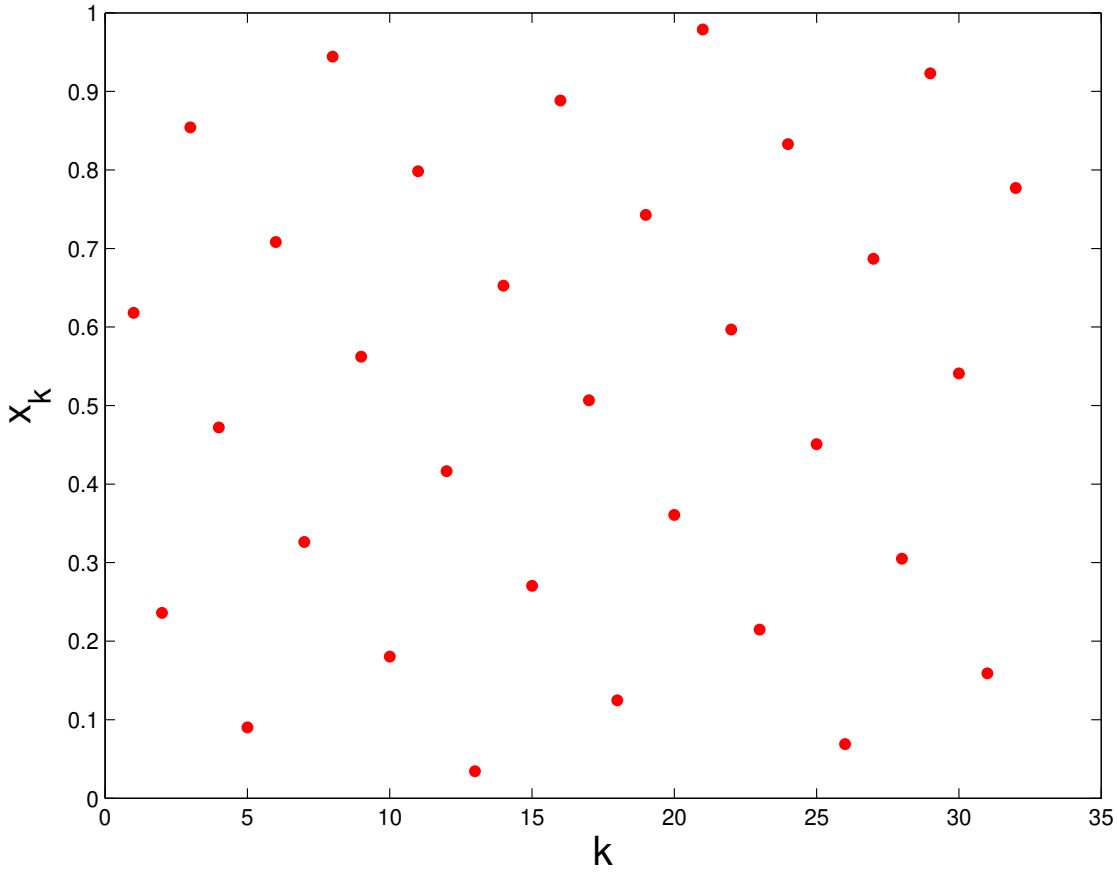
For instance, $z = \varphi = (\sqrt{5}+1)/2 \simeq 1.618034 =$ Golden section

first 100 points x_k



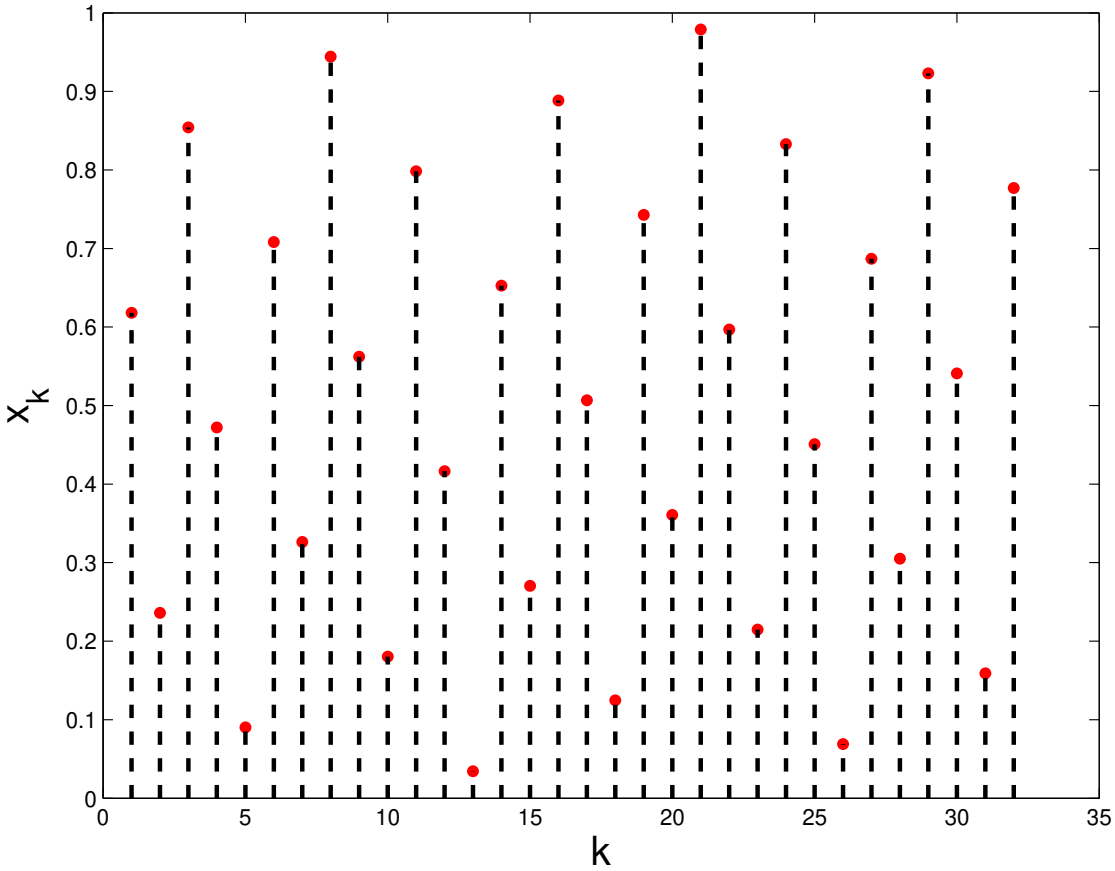
For $d > 1$:

$$x_k = \{k\varphi\},\ k = 1, \ldots, 32$$



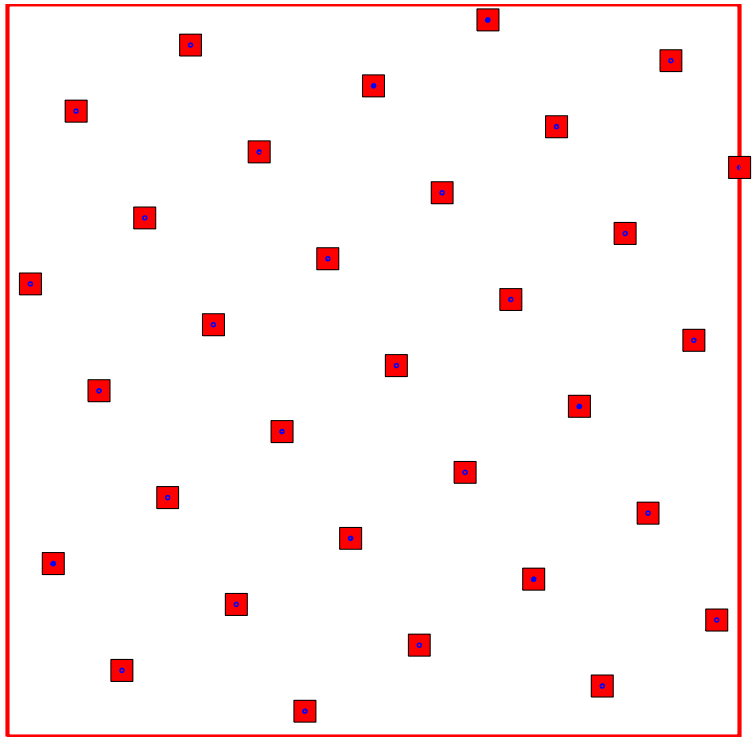
For $d > 1$:

$$x_k = \{k\varphi\},\ k = 1, \ldots, 32$$



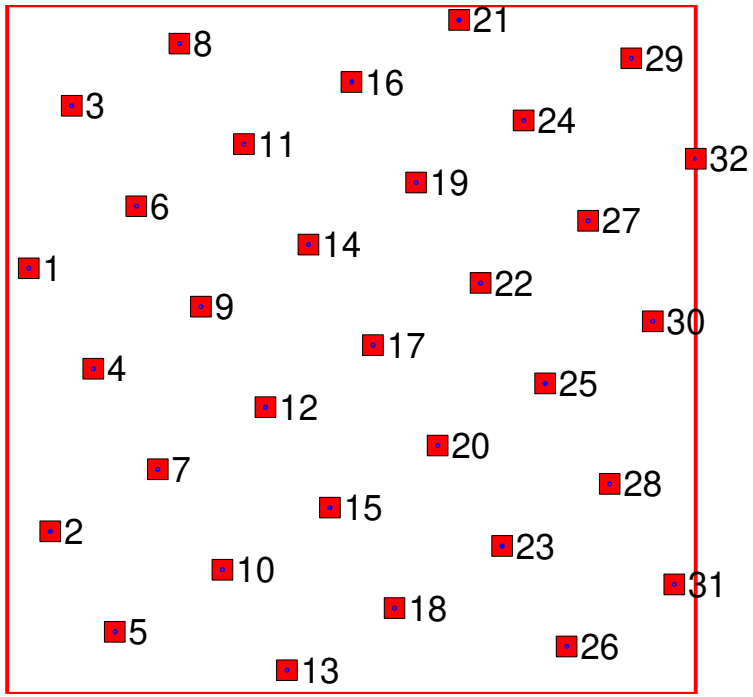
For $d > 1$:

$$\mathbf{x}_k = (k/n, \{k\varphi\})^\top$$



For $d > 1$:

$$\mathbf{x}_k = (k/n, \{k\varphi\})^\top$$



➠ Replace k/n (monotonically increasing and only valid for $k = 1, \ldots, n$) by $\{kz\}$, z irrational $\rightarrow \mathbf{x}_k \in [0,1]^2$, $k = 1, 2, 3 \ldots$

➠ repeat... $\rightarrow$ recursively $\mathbf{x}_k \in [0,1]^d$, for any given d

Take $\mathbf{x}_k \triangleq \{k\mathbf{z}\}$, $\mathbf{z}$ irrational vector in $\mathbb{R}^d$
(with independent components over rationals: $\mathbf{q}^\top \mathbf{z} \neq 0$, $\forall \mathbf{q} \in \mathbb{Q}^d$)

- sequence $\mathbf{X}_\infty$ uniformly distributed in $[0,1]^d$

  (Kuipers and Niederreiter, 1974, p. 48)

- $\forall \epsilon > 0$, $D_n(\mathbf{X}_\infty) = \mathcal{O}\left(\frac{(1+\log n)^{d+1+\epsilon}}{n}\right)$ for almost all $\mathbf{z}$

Interesting, but does not say which $\mathbf{z}$ we should take

Take $\mathbf{x}_k \triangleq \{k\mathbf{z}\}$, $\mathbf{z}$ irrational vector in $\mathbb{R}^d$
(with independent components over rationals: $\mathbf{q}^\top \mathbf{z} \neq 0$, $\forall \mathbf{q} \in \mathbb{Q}^d$)

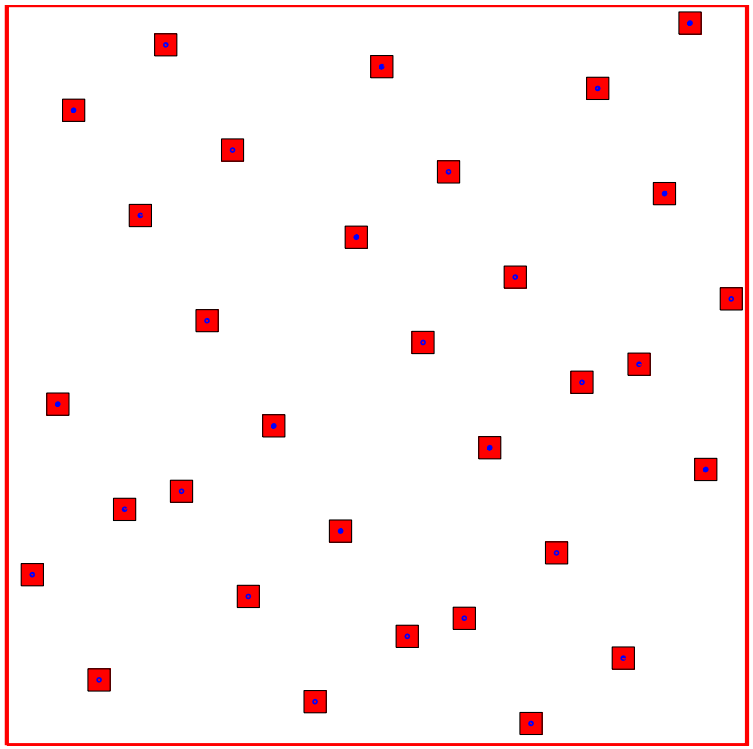
- sequence $\mathbf{X}_\infty$ uniformly distributed in $[0,1]^d$

  (Kuipers and Niederreiter, 1974, p. 48)

- $\forall \epsilon > 0$, $D_n(\mathbf{X}_\infty) = \mathcal{O}\left(\frac{(1+\log n)^{d+1+\epsilon}}{n}\right)$ for almost all $\mathbf{z}$

Interesting, but does not say which $\mathbf{z}$ we should take

$d = 2$, $n = 32$, $\mathbf{z} = (\varphi, (\sqrt{2}-1)^2)^\top$



Take $\mathbf{x}_k \triangleq \{k\mathbf{z}\}$, $\mathbf{z}$ irrational vector in $\mathbb{R}^d$
(with independent components over rationals: $\mathbf{q}^\top \mathbf{z} \neq 0$, $\forall \mathbf{q} \in \mathbb{Q}^d$)

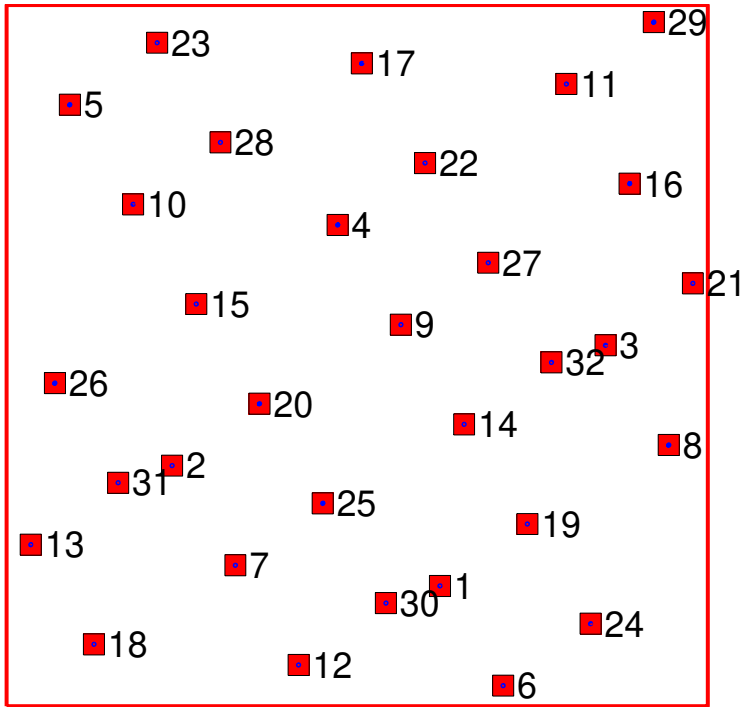
➠ sequence $\mathbf{X}_\infty$ uniformly distributed in $[0,1]^d$

(Kuipers and Niederreiter, 1974, p. 48)

➠ $\forall \epsilon > 0$, $D_n(\mathbf{X}_\infty) = \mathcal{O}\left(\frac{(1+\log n)^{d+1+\epsilon}}{n}\right)$ for almost all $\mathbf{z}$

Interesting, but does not say which $\mathbf{z}$ we should take

$d = 2$, $n = 32$, $\mathbf{z} = (\varphi, (\sqrt{2}-1)^2)^\top$



Lattices: $\mathbf{z} = \mathbf{g}/n$, with $\mathbf{g} \in \mathbb{Z}^d$

- $\mathbf{x}_k \triangleq \left\{ \frac{k}{n} \mathbf{g} \right\}$ (with repetitions if $k \geq n$)
  - n points $\neq$ if $\gcd(g_1, \ldots, g_d, n) = 1$
  - n points $\neq$ for each coordinate if $\gcd(g_i, n) = 1$ for all i
- Regular arrangement of points ($\approx$ grid)
  - Regularity of $f(\cdot)$ may be accounted for in integration error bounds

Lattices: $\mathbf{z} = \mathbf{g}/n$, with $\mathbf{g} \in \mathbb{Z}^d$

- $\mathbf{x}_k \triangleq \left\{ \frac{k}{n} \mathbf{g} \right\}$ (with repetitions if $k \geq n$)
  - n points $\neq$ if $\gcd(g_1, \ldots, g_d, n) = 1$
  - n points $\neq$ for each coordinate if $\gcd(g_i, n) = 1$ for all i

- Regular arrangement of points ($\approx$ grid)
  - Regularity of $f(\cdot)$ may be accounted for in integration error bounds

For $d = 2$, $\mathbf{g} = (1, \ F_{m-1})^\top$ for $n = F_m$ is a very good choice,
with (F_m) = Fibonacci sequence: $F_1 = F_2 = 1$, $F_{k+1} = F_k + F_{k-1}$, $k \geq 2$

Since $F_{m-1}/F_m \to 1/\varphi$ for $m \to \infty$, the construction is similar to
$\mathbf{x}_k = (\frac{k}{n}, \ \left\{ \frac{k}{\varphi} \right\})^\top = (\frac{k}{n}, \ \{k\varphi\})^\top$, $k = 1, \ldots, n$

➤ Strong connection with optimal design for Fourier regression (sin, cos) (Bates et al., 1996; Riccomagno et al., 1997)

➤ The exist constructions (non explicit) with good properties (good lattice points) ➠ tables (Maisonneuve, 1972)

➤ Korobov (1960) suggests $\mathbf{g} = (1, g, g^2, \ldots, g^{d-1})^\top$

➤ Strong connection with optimal design for Fourier regression (sin, cos) (Bates et al., 1996; Riccomagno et al., 1997)

➤ The exist constructions (non explicit) with good properties (good lattice points) ➠ tables (Maisonneuve, 1972)

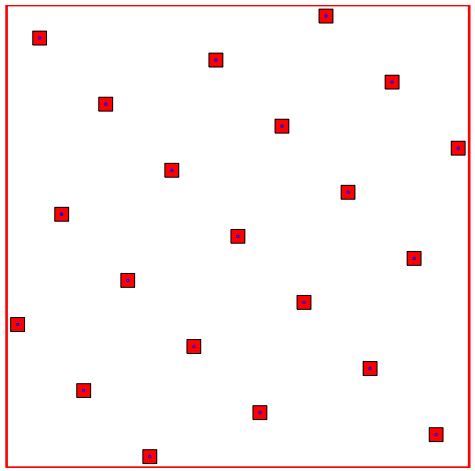
➤ Korobov (1960) suggests $\mathbf{g} = (1, g, g^2, \ldots, g^{d-1})^\top$

➤ Optimization of $\mathbf{g}$ in, e.g., (Sloan and Walsh, 1990; Sloan and Reztsov, 2002; Nuyens, 2007)

Ex: $d = 2$, $n = 21 = F_8$

$$\mathbf{g} = (1, F_7)$$

$\Phi_{Mm} = 0.2020$
$\Phi_{mM} = 0.2357$
$D_{Cent,L_2} = 0.0280$
$D_{WA,L_2} = 0.0388$



➤ Strong connection with optimal design for Fourier regression (sin, cos) (Bates et al., 1996; Riccomagno et al., 1997)

➤ The exist constructions (non explicit) with good properties (good lattice points) ➠ tables (Maisonneuve, 1972)

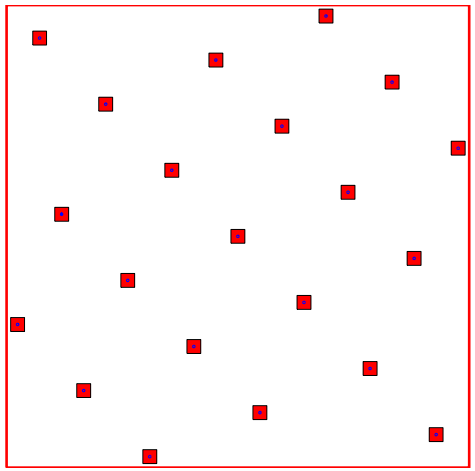
➤ Korobov (1960) suggests $\mathbf{g} = (1, g, g^2, \ldots, g^{d-1})^\top$

➤ Optimization of $\mathbf{g}$ in, e.g., (Sloan and Walsh, 1990; Sloan and Reztsov, 2002; Nuyens, 2007)

Ex: $d = 2$, $n = 21 = F_8$

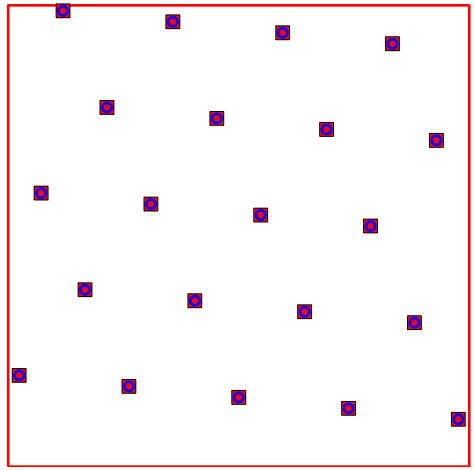
$\mathbf{g} = (1, F_7)$

$\Phi_{Mm} = 0.2020$
$\Phi_{mM} = 0.2357$
$D_{Cent,L_2} = 0.0280$
$D_{WA,L_2} = 0.0388$



$\Phi_{Mm} = 0.2302$
$\Phi_{mM} = 0.2217$
$D_{Cent,L_2} = 0.0536$
$D_{WA,L_2} = 0.0633$

$\mathbf{g} = (1, g^*)$
(optimal for Φ_{Mm})



- Can only generate n points:
  - infinite sequence in $[0,1]^d$ if $\mathbf{x}_k \triangleq \{u_k\,\mathbf{g}\}$

  with (u_k) a (scalar) LDS (Hickernell, 1998b)

- Can only generate n points:
  - infinite sequence in $[0,1]^d$ if $\mathbf{x}_k \triangleq \{u_k\,\mathbf{g}\}$
    with (u_k) a (scalar) LDS (Hickernell, 1998b)

We only considered the rank-on rule, there also exist

rank r rule: $\mathbf{x}_{k_1,\ldots,k_r} \triangleq \left\{ \frac{k_1}{n_1}\mathbf{g}_1 + \frac{k_2}{n_2}\mathbf{g}_2 + \cdots + \frac{k_r}{n_r}\mathbf{g}_r \right\}$, $k_j \in \{1,\ldots,n_j\}$

Copy rule: divide $[0,1]^d$ into k^d cubes with edge length $1/k$, construct a lattice in each

van der Corput (1935) LDS

$\boxed{d = 1}$

Let $\mathcal{Z}_b = \{0, 1, \ldots, b-1\}$ be the alphabet for base $b \geq 2$
(e.g., $\mathcal{Z}_2 = \{0, 1\}$, $\mathcal{Z}_3 = \{0, 1, 2\}$)

- Any $k = 0, 1, \ldots, b^m - 1$ can be written as $k = \sum_{\ell=0}^{m-1} a_\ell b^\ell$
with m characters $a_0, a_1, \ldots, a_{m-1}$ (dependent on k)
(that is, $k = \underline{a_{m-1} a_{m-2} \cdots a_2 a_1 a_0}_b$)

- To k, we associate $\Phi_b(k) = \sum_{\ell=0}^{m-1} a_\ell b^{-(\ell+1)}$
  - The van der Corput sequence in base b is defined by $\boxed{x_k \triangleq \Phi_b(k)}$

base $b = 2$ (van der Corput, 1935)

k	k in base 2	$\phi_b(k)$ in base 2	$\phi_2(k)$
0	0	0.0	0
1	1	0.1	1/2
2	10	0.01	1/4
3	11	0.11	3/4
4	100	0.001	1/8
5	101	0.101	5/8
k	$\underline{a_{m-1}\cdots a_1 a_0}_2$	$0.a_0 a_1 \cdots a_{m-1}$	$\sum_{\ell=0}^{m-1} a_\ell 2^{-(\ell+1)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

$$\Rightarrow \quad nD_n^*(\mathbf{X}_\infty) = nD_n(\mathbf{X}_\infty) \leq 1 + \frac{\log n}{\log 8}$$

base $b = 3$

k	k in base 3	$\phi_3(k)$
0	0	0
1	1	1/3
2	2	2/3
3	10	1/9
4	11	4/9
5	12	7/9
6	20	2/9
7	21	5/9
8	22	8/9
⋮	⋮	⋮

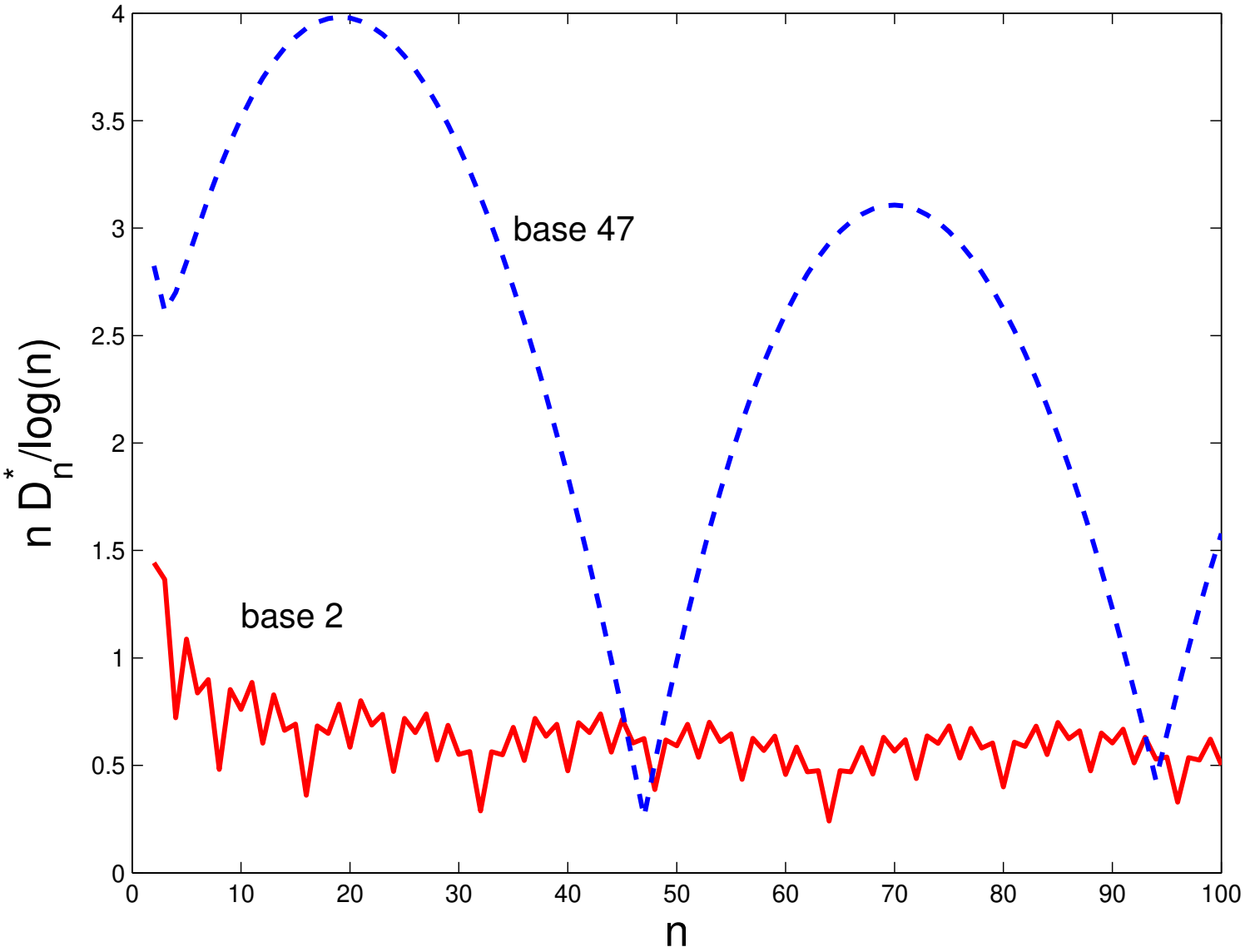
base $b = 3$

k	k in base 3	$\phi_3(k)$
0	0	0
1	1	1/3
2	2	2/3
3	10	1/9
4	11	4/9
5	12	7/9
6	20	2/9
7	21	5/9
8	22	8/9
⋮	⋮	⋮

Particular choice of b + suitable permutation of $\mathcal{Z}_b$

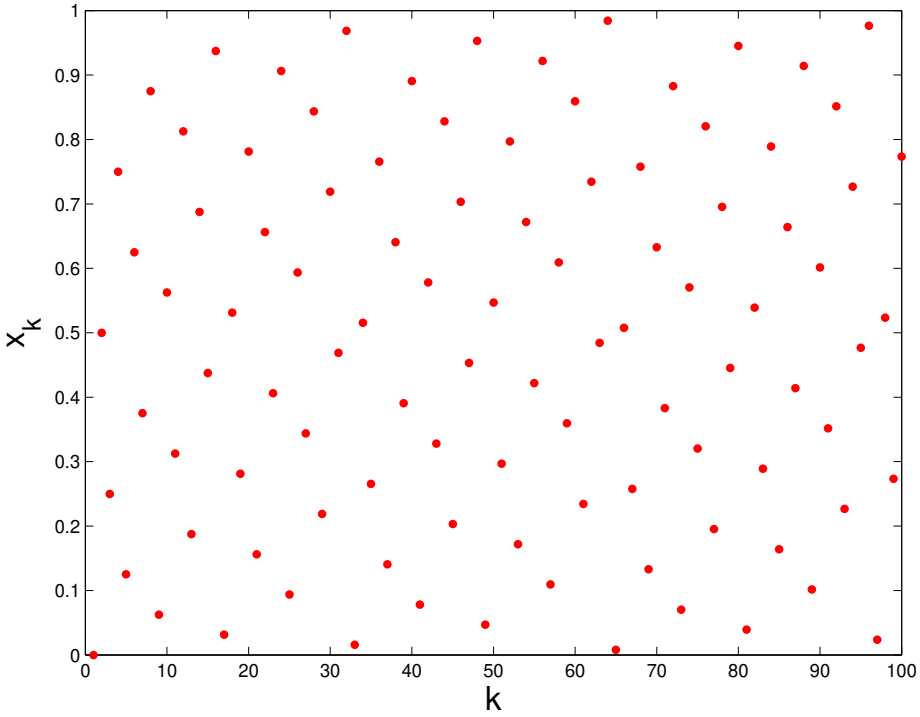
➠ best known performance for $\limsup_{n\to\infty} nD_n^*/log(n)$ $(b = 12)$ and $\limsup_{n\to\infty} nD_n/log(n)$ $(b = 36)$ (results by H. Faure (1977–20xx))

van der Corput : $(n/\log n) \times D_n^*$ for $b = 2$ and $b = 47$

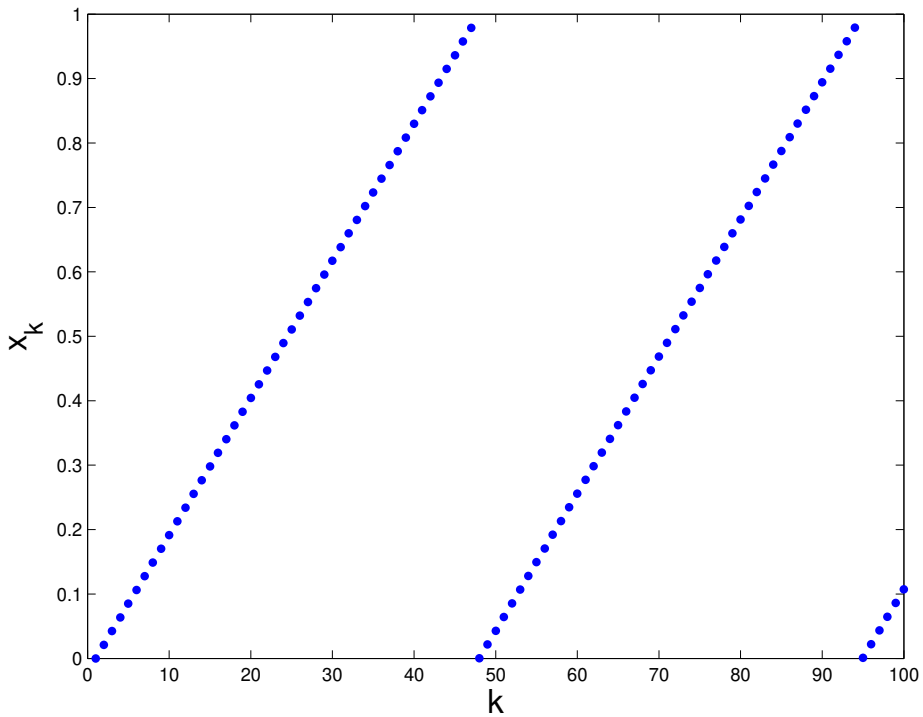


van der Corput : $x_1, x_2 \ldots, x_{100}$

$b = 2$



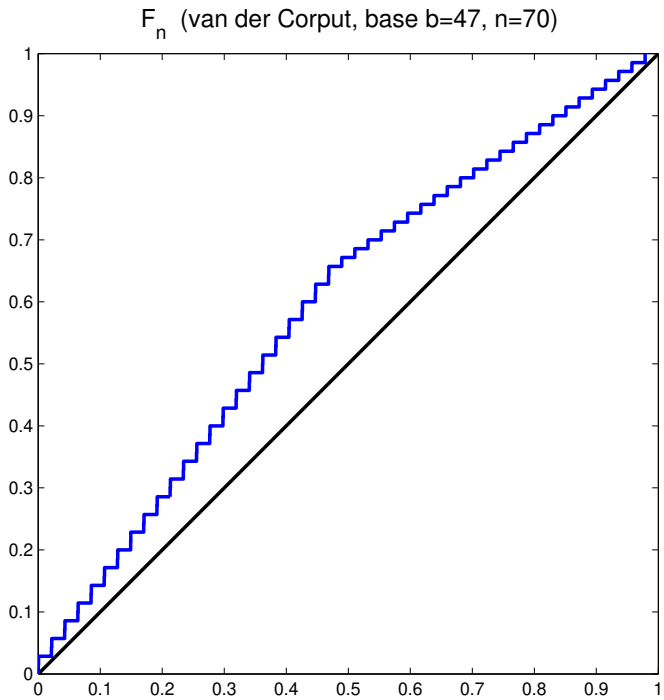
$b = 47$, $x_1, x_2 \ldots$



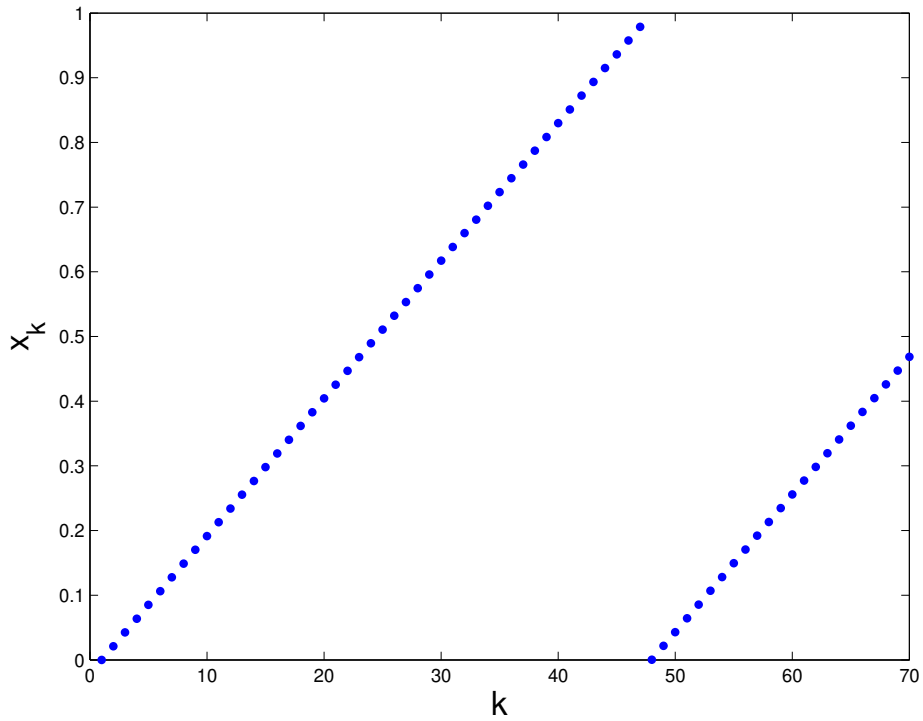
. . . and, moreover, $[0, 1]$ is filled in a particular order

van der Corput : $x_1, x_2 \ldots, x_{70}$

$b = 47$, empirical c.d.f.



$b = 47$, $x_1, x_2 \ldots$



. . . and, moreover, $[0, 1]$ is filled in a particular order

$\boxed{d > 1}$: work separately on each component

$b_1, b_2 \ldots$ integers such that $\gcd(b_i, b_j) = 1$ for all $i \neq j$

(in practice, first prime numbers:

$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47 \ldots$)

$\boxed{d > 1}$: work separately on each component

$b_1, b_2 \ldots$ integers such that $\gcd(b_i, b_j) = 1$ for all $i \neq j$

(in practice, first prime numbers:

$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47 \ldots$)

- Hammersley: $\mathbf{x}_k \triangleq (k/n, \phi_{b_1}(k), \ldots, \phi_{b_{d-1}}(k))^\top$, $k = 1, \ldots, n$

  (n is given — fixed)

  $$D_n^*(\mathbf{X}_n) \leq A_{d-1} \frac{(\log n)^{d-1}}{n} + \mathcal{O}\left(\frac{(\log n)^{d-2}}{n}\right)$$

$\boxed{d > 1}$: work separately on each component

$b_1, b_2 \ldots$ integers such that $\gcd(b_i, b_j) = 1$ for all $i \neq j$

(in practice, first prime numbers:

$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47 \ldots$)

- Hammersley: $\mathbf{x}_k \triangleq (k/n, \phi_{b_1}(k), \ldots, \phi_{b_{d-1}}(k))^\top$, $k = 1, \ldots, n$

  (n is given — fixed)

  $$D_n^*(\mathbf{X}_n) \leq A_{d-1} \frac{(\log n)^{d-1}}{n} + \mathcal{O}\left(\frac{(\log n)^{d-2}}{n}\right)$$

- Halton: $\mathbf{x}_k \triangleq (\phi_{b_1}(k), \ldots, \phi_{b_{d-1}}, \phi_{b_d}(k))^\top$, $k = 1, 2, 3 \ldots$

  $$D_n^*(\mathbf{X}_\infty) \leq A_d \frac{(\log n)^d}{n} + \mathcal{O}\left(\frac{(\log n)^{d-1}}{n}\right)$$

$\boxed{d > 1}$: work separately on each component

$b_1, b_2 \ldots$ integers such that $\gcd(b_i, b_j) = 1$ for all $i \neq j$

(in practice, first prime numbers:

$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47 \ldots$)

- Hammersley: $\mathbf{x}_k \triangleq (k/n, \phi_{b_1}(k), \ldots, \phi_{b_{d-1}}(k))^\top$, $k = 1, \ldots, n$

  (n is given — fixed)

  $$D_n^*(\mathbf{X}_n) \leq A_{d-1} \frac{(\log n)^{d-1}}{n} + \mathcal{O}\left(\frac{(\log n)^{d-2}}{n}\right)$$

- Halton: $\mathbf{x}_k \triangleq (\phi_{b_1}(k), \ldots, \phi_{b_{d-1}}, \phi_{b_d}(k))^\top$, $k = 1, 2, 3 \ldots$

  $$D_n^*(\mathbf{X}_\infty) \leq A_d \frac{(\log n)^d}{n} + \mathcal{O}\left(\frac{(\log n)^{d-1}}{n}\right)$$

It is conjectured that (only proved for $d = 1, 2$)

$$D_n^*(\mathbf{X}_n) \geq B_d \frac{(\log n)^{d-1}}{n} \text{ and}$$

$$D_n^*(\mathbf{X}_\infty) \geq B'_d \frac{(\log n)^d}{n} \text{ infinitely often}$$

The speed of decrease of D_n^* is thus optimal for Hammersley and Halton

$\boxed{d > 1}$: work separately on each component

$b_1, b_2 \ldots$ integers such that $\gcd(b_i, b_j) = 1$ for all $i \neq j$

(in practice, first prime numbers:

$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47 \ldots$)

- Hammersley: $\mathbf{x}_k \triangleq (k/n, \phi_{b_1}(k), \ldots, \phi_{b_{d-1}}(k))^\top$, $k = 1, \ldots, n$

  (n is given — fixed)

  $$D_n^*(\mathbf{X}_n) \leq A_{d-1} \frac{(\log n)^{d-1}}{n} + \mathcal{O}\left(\frac{(\log n)^{d-2}}{n}\right)$$

- Halton: $\mathbf{x}_k \triangleq (\phi_{b_1}(k), \ldots, \phi_{b_{d-1}}, \phi_{b_d}(k))^\top$, $k = 1, 2, 3 \ldots$

  $$D_n^*(\mathbf{X}_\infty) \leq A_d \frac{(\log n)^{d}}{n} + \mathcal{O}\left(\frac{(\log n)^{d-1}}{n}\right)$$

It is conjectured that (only proved for $d = 1, 2$)

$$D_n^*(\mathbf{X}_n) \geq B_d \frac{(\log n)^{d-1}}{n} \text{ and}$$

$$D_n^*(\mathbf{X}_\infty) \geq B'_d \frac{(\log n)^{d}}{n} \text{ infinitely often}$$

The speed of decrease of D_n^* is thus optimal for Hammersley and Halton

. . . but what about the constant A_d?

It can be shown that $\lim_{d\to\infty} \frac{\log A_d}{d \log d} = 1$

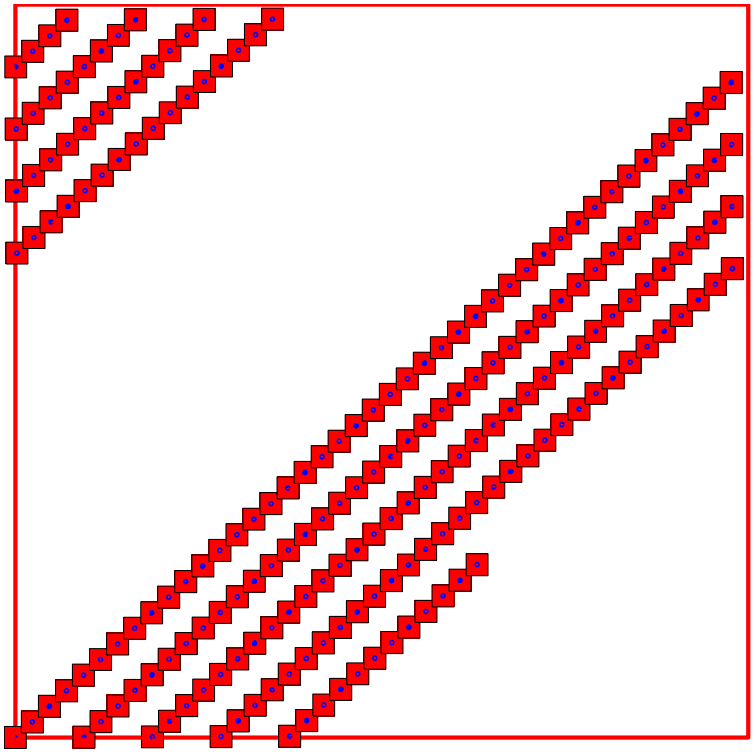
($A_d \nearrow$ super-exponentially fast with d!)

It can be shown that $\lim_{d\to\infty} \frac{\log A_d}{d \log d} = 1$

($A_d \nearrow$ super-exponentially fast with $d!$)

➠ Discrepancy is not very good for large d

Halton LDS, $d = 15$ ($\Rightarrow b_{15} = 47$) :
$\{\mathbf{x}_k\}_{14}$ and $\{\mathbf{x}_k\}_{15}$, $k = 1, \ldots, 200$

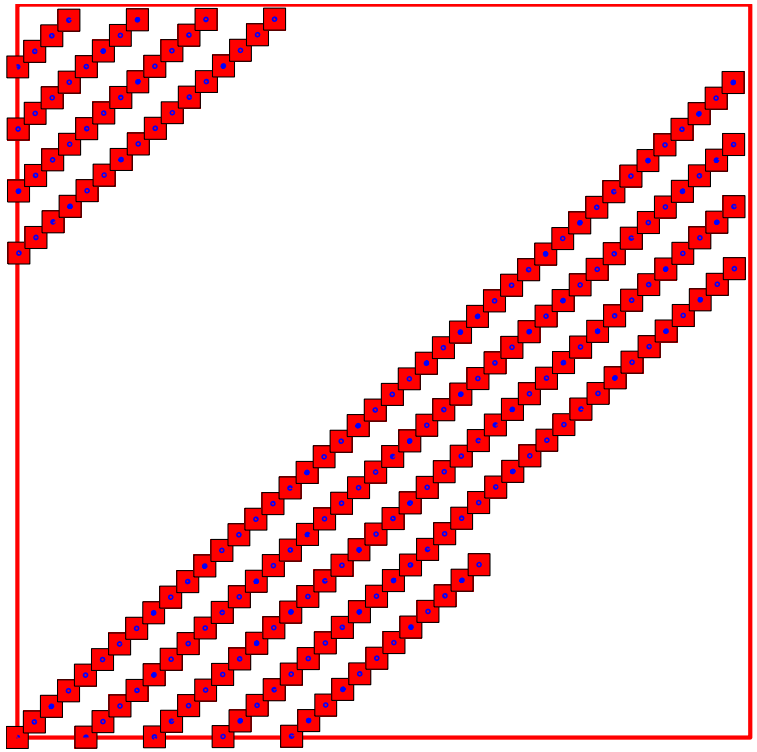


It can be shown that $\lim_{d\to\infty} \frac{\log A_d}{d \log d} = 1$

($A_d \nearrow$ super-exponentially fast with d!)

➠ Discrepancy is not very good for large d

Halton LDS, $d = 15$ ($\Rightarrow b_{15} = 47$) :
$\{\mathbf{x}_k\}_{14}$ and $\{\mathbf{x}_k\}_{15}$, $k = 1, \ldots, 200$



. . . not too bad in the plane $\{\mathbf{x}_k\}_{d_1}, \{\mathbf{x}_k\}_{d_2}$ if $n = b_{d_1} b_{d_2}$
(with $b_d \approx d(\log d + \log\log d)$)

2.5 (t, m, d)-nets & (t, d)-sequences (Niederreiter, 1992, Chap. 4), (Owen, 1995)

suites-(t, d)

2.5 (t, m, d)-nets & (t, d)-sequences (Niederreiter, 1992, Chap. 4), (Owen, 1995)

suites-(t, d)

Motivation: overcome the issue $A_d \nearrow \infty$ as $d \to \infty$ in Halton LDS

For a base b, consider an elementary interval (= a d-dimensional box)

$$\mathbb{P}(\mathbf{a}, \mathbf{q}) = \prod_{j=1}^{d} [\frac{a_j}{b^{q_j}}, \frac{1+a_j}{b^{q_j}}]$$

where q_j and a_j are integers, $0 \leq q_j$ and $0 \leq a_j \leq b^{q_j} - 1$

$$\mathbb{P}(\mathbf{a}, \mathbf{q}) \subset [0,1]^d \text{ and } \text{vol}[\mathbb{P}(\mathbf{a}, \mathbf{q})] = \prod_{j=1}^{d} b^{-q_j} = b^{-\sum_{j=1}^{d} q_j}$$

2.5 (t, m, d)-nets & (t, d)-sequences (Niederreiter, 1992, Chap. 4), (Owen, 1995)

suites-(t, d)

Motivation: overcome the issue $A_d \nearrow \infty$ as $d \to \infty$ in Halton LDS

For a base b, consider an elementary interval (= a d-dimensional box)

$$\mathbb{P}(\mathbf{a}, \mathbf{q}) = \prod_{j=1}^{d} [\frac{a_j}{b^{q_j}}, \frac{1+a_j}{b^{q_j}}]$$

where q_j and a_j are integers, $0 \leq q_j$ and $0 \leq a_j \leq b^{q_j} - 1$

$$\mathbb{P}(\mathbf{a}, \mathbf{q}) \subset [0,1]^d \text{ and } \text{vol}[\mathbb{P}(\mathbf{a}, \mathbf{q})] = \prod_{j=1}^{d} b^{-q_j} = b^{-\sum_{j=1}^{d} q_j}$$

Objective: put points in each elementary interval
(considering all possible cuts into elementary intervals)

2.5 (t, m, d)-nets & (t, d)-sequences (Niederreiter, 1992, Chap. 4), (Owen, 1995)

suites-(t, d)

Motivation: overcome the issue $A_d \nearrow \infty$ as $d \to \infty$ in Halton LDS

For a base b, consider an elementary interval (= a d-dimensional box)

$$\mathbb{P}(\mathbf{a}, \mathbf{q}) = \prod_{j=1}^{d} [\frac{a_j}{b^{q_j}}, \frac{1+a_j}{b^{q_j}}]$$

where q_j and a_j are integers, $0 \leq q_j$ and $0 \leq a_j \leq b^{q_j} - 1$

$$\mathbb{P}(\mathbf{a}, \mathbf{q}) \subset [0,1]^d \text{ and } \text{vol}[\mathbb{P}(\mathbf{a}, \mathbf{q})] = \prod_{j=1}^{d} b^{-q_j} = b^{-\sum_{j=1}^{d} q_j}$$

Objective: put points in each elementary interval
(considering all possible cuts into elementary intervals)

More precisely: for $0 \leq t \leq m$, a $\boxed{(t, m, d)\text{-net}}$ in base b contains $n = b^m$ points, such that each elementary interval with volume b^{t-m} contains b^t points

Example: $(0, 2, 2)$-net in base 2 ($b = 2$, $d = 2$, $m = 2$, $t = 0$)

➠ $n = b^m = 4$, $b^0 = 1$ point in each elementary interval with volume $b^{t-m} = 1/4$

Example: $(0, 2, 2)$-net in base 2 ($b = 2$, $d = 2$, $m = 2$, $t = 0$)

➠ $n = b^m = 4$, $b^0 = 1$ point in each elementary interval with volume $b^{t-m} = 1/4$

$$\sum_{j=1}^{d} q_j = m - t = 2 \Rightarrow q_j \in \{0, 1, 2\}$$

	q_1	q_2	
(i)	0	2	$a_1 = 0$, $a_2 \in \{0, 1, 2, 3\}$
(ii)	2	0	$a_1 \in \{0, 1, 2, 3\}$, $a_2 = 0$
(iii)	1	1	$a_1 \in \{0, 1\}$, $a_2 \in \{0, 1\}$

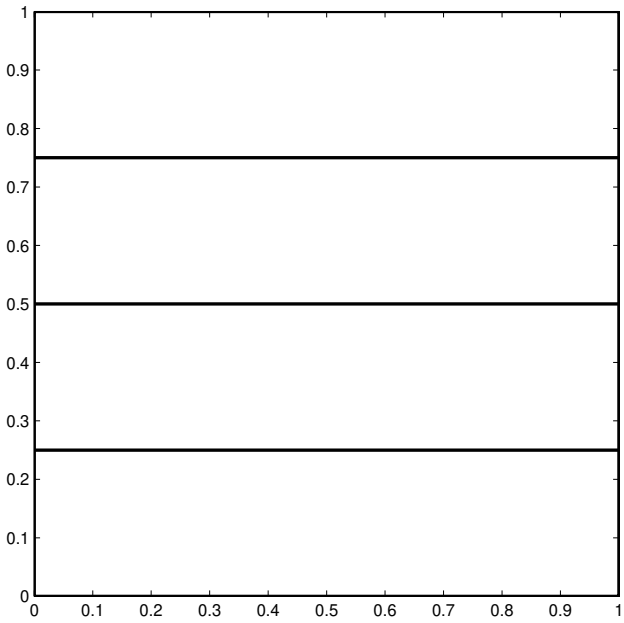
Example: $(0, 2, 2)$-net in base 2 ($b = 2$, $d = 2$, $m = 2$, $t = 0$)

➠ $n = b^m = 4$, $b^0 = 1$ point in each elementary interval with volume $b^{t-m} = 1/4$

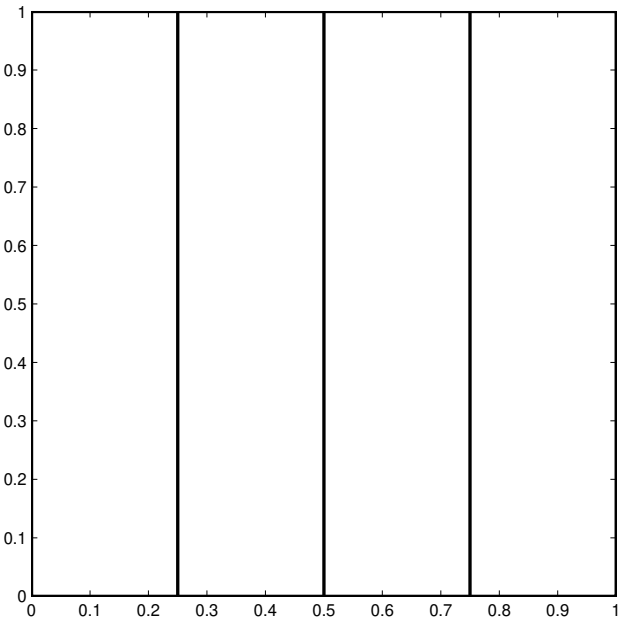
$$\sum_{j=1}^{d} q_j = m - t = 2 \Rightarrow q_j \in \{0, 1, 2\}$$

	q_1	q_2	
(i)	0	2	$a_1 = 0$, $a_2 \in \{0, 1, 2, 3\}$
(ii)	2	0	$a_1 \in \{0, 1, 2, 3\}$, $a_2 = 0$
(iii)	1	1	$a_1 \in \{0, 1\}$, $a_2 \in \{0, 1\}$

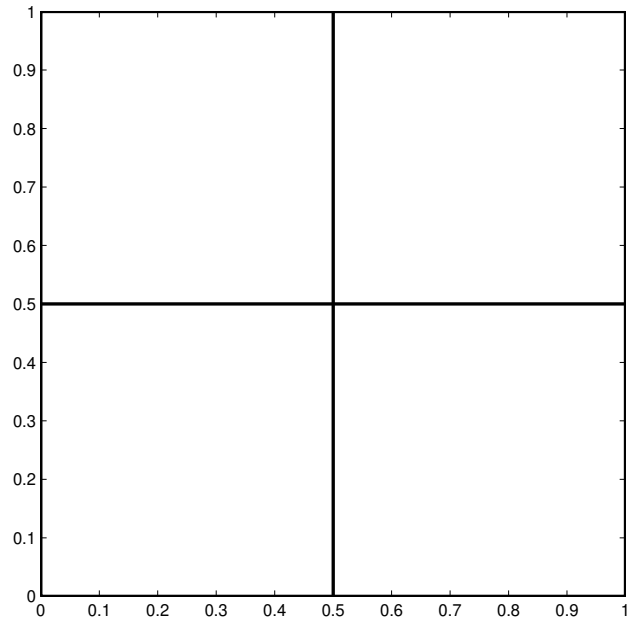
case (i)



case (ii)



case (iii)



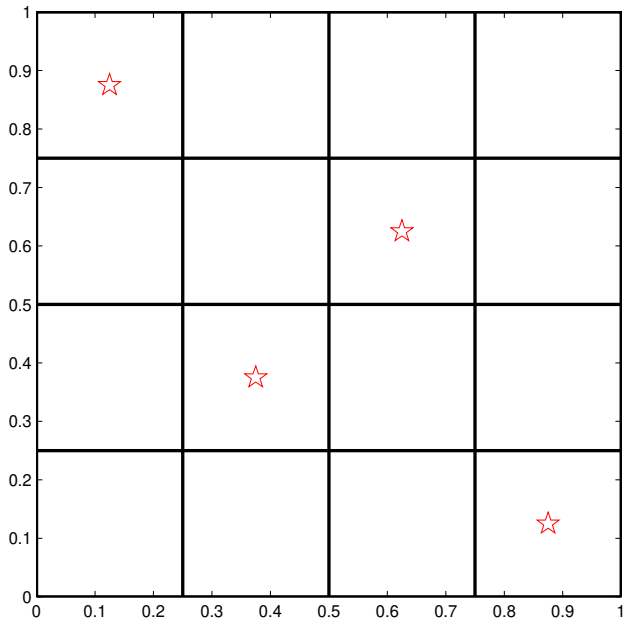
Example: $(0, 2, 2)$-net in base 2 ($b = 2$, $d = 2$, $m = 2$, $t = 0$)

➠ $n = b^m = 4$, $b^0 = 1$ point in each elementary interval with volume $b^{t-m} = 1/4$

$$\sum_{j=1}^{d} q_j = m - t = 2 \Rightarrow q_j \in \{0, 1, 2\}$$

	q_1	q_2	
(i)	0	2	$a_1 = 0$, $a_2 \in \{0, 1, 2, 3\}$
(ii)	2	0	$a_1 \in \{0, 1, 2, 3\}$, $a_2 = 0$
(iii)	1	1	$a_1 \in \{0, 1\}$, $a_2 \in \{0, 1\}$

a $(0, 2, 2)$-net in base 2



A few properties :

- $t = m$: trivial situation ➠ $b^t = b^m = n$ points (all) in the interval with volume $b^{t-m} = 1$

A few properties :

- $t = m$: trivial situation ➠ $b^t = b^m = n$ points (all) in the interval with volume $b^{t-m} = 1$
- (t, m, d)-net in base $b \Rightarrow (t', m, d)$-net in base b for $t \leq t' \leq m$
  ➠ difficulty increases as t decreases

A few properties :

- $t = m$: trivial situation ➠ $b^t = b^m = n$ points (all) in the interval with volume $b^{t-m} = 1$
- (t, m, d)-net in base $b \Rightarrow (t', m, d)$-net in base b for $t \leq t' \leq m$
  ➠ difficulty increases as t decreases
- $t = 0$, $m = 1$, base b
  $\Rightarrow n = b^m = b$ points, $b^t = 1$ point per elementary interval, $\sum_{j=1}^{d} q_j = 1 \Rightarrow$ a unique $q_j \neq 0$, $q_j = 1 \Rightarrow a_j \in \{0, 1, \ldots, b-1\}$
  ➠ a $(0, 1, d)$-net in base b is a Lh with b points

A few properties :

- $t = m$: trivial situation ➠ $b^t = b^m = n$ points (all) in the interval with volume $b^{t-m} = 1$
- (t, m, d)-net in base $b \Rightarrow (t', m, d)$-net in base b for $t \leq t' \leq m$
  ➠ difficulty increases as t decreases
- $t = 0$, $m = 1$, base b
  $\Rightarrow n = b^m = b$ points, $b^t = 1$ point per elementary interval, $\sum_{j=1}^{d} q_j = 1 \Rightarrow$ a unique $q_j \neq 0$, $q_j = 1 \Rightarrow a_j \in \{0, 1, \ldots, b-1\}$
  ➠ a $(0, 1, d)$-net in base b is a Lh with b points
- A $(0, 2, d)$-net in base b is an Orthogonal Array (OA) and a Lh with $n = b^2$ points: previous example had $b = 2$ and $d = 2$
  — but the construction is not always possible...
- no $(0, m, d)$-net in base b for $d > b + 1$ (Niederreiter, 1992, p. 62)

$\mathbf{X}_\infty =$ $\boxed{(t,d)\text{-sequence}}$ in base b if for any $k \geq 0$ and any $m \geq t$ the b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \ldots, \mathbf{x}_{(k+1)b^m-1}\}$ form a (t, m, d)-net

$\mathbf{X}_\infty =$ (t,d)-sequence in base b if for any $k \geq 0$ and any $m \geq t$ the b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \ldots, \mathbf{x}_{(k+1)b^m-1}\}$ form a (t, m, d)-net

A few properties:

- Especially interesting for t small

$\mathbf{X}_\infty =$ (t, d)-sequence in base b if for any $k \geq 0$ and any $m \geq t$ the b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \ldots, \mathbf{x}_{(k+1)b^m-1}\}$ form a (t, m, d)-net

A few properties:

- Especially interesting for t small
- no $(0, d)$-sequence in base b for $d > b$ (Niederreiter, 1992, p. 62)
  $(0, d)$-sequence in base $b \geq d$ with b prime = Faure (1982) sequence

$\mathbf{X}_\infty =$ (t,d)-sequence in base b if for any $k \geq 0$ and any $m \geq t$ the b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \ldots, \mathbf{x}_{(k+1)b^m-1}\}$ form a (t, m, d)-net

A few properties:

- Especially interesting for t small
- no $(0, d)$-sequence in base b for $d > b$ (Niederreiter, 1992, p. 62)
  $(0, d)$-sequence in base $b \geq d$ with b prime = Faure (1982) sequence
- van der Corput sequence in base b is a $(0, 1)$-sequence in base b

$\mathbf{X}_\infty =$ (t,d)-sequence in base b if for any $k \geq 0$ and any $m \geq t$ the b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \ldots, \mathbf{x}_{(k+1)b^m-1}\}$ form a (t, m, d)-net

A few properties:

- Especially interesting for t small
- no $(0, d)$-sequence in base b for $d > b$ (Niederreiter, 1992, p. 62)
  $(0, d)$-sequence in base $b \geq d$ with b prime = Faure (1982) sequence
- van der Corput sequence in base b is a $(0, 1)$-sequence in base b
- Construction rather complicated...
  For $b = 2$, any d: Sobol' (1967) sequences, with a smaller t when $d \geq 8$ for Niederreiter (1992) sequences

(t, d)-sequences in base b rely on rather complicated algebraic constructions, but ensure a good distribution of points in $\mathbb{I}_d = [0, 1]^d$

- t and b should be as small as possible
- Sobol' (1967) sequences: $b = 2$
- Niederreiter (1992) sequences: $b = 2$ and t smaller than for Sobol' when $d \geq 8$
- Faure (1982) $(0, d)$-sequences: base $b \geq d$, with b prime

(t, d)-sequences in base b rely on rather complicated algebraic constructions, but ensure a good distribution of points in $\mathbb{I}_d = [0, 1]^d$

- t and b should be as small as possible
- Sobol' (1967) sequences: $b = 2$
- Niederreiter (1992) sequences: $b = 2$ and t smaller than for Sobol' when $d \geq 8$
- Faure (1982) $(0, d)$-sequences: base $b \geq d$, with b prime

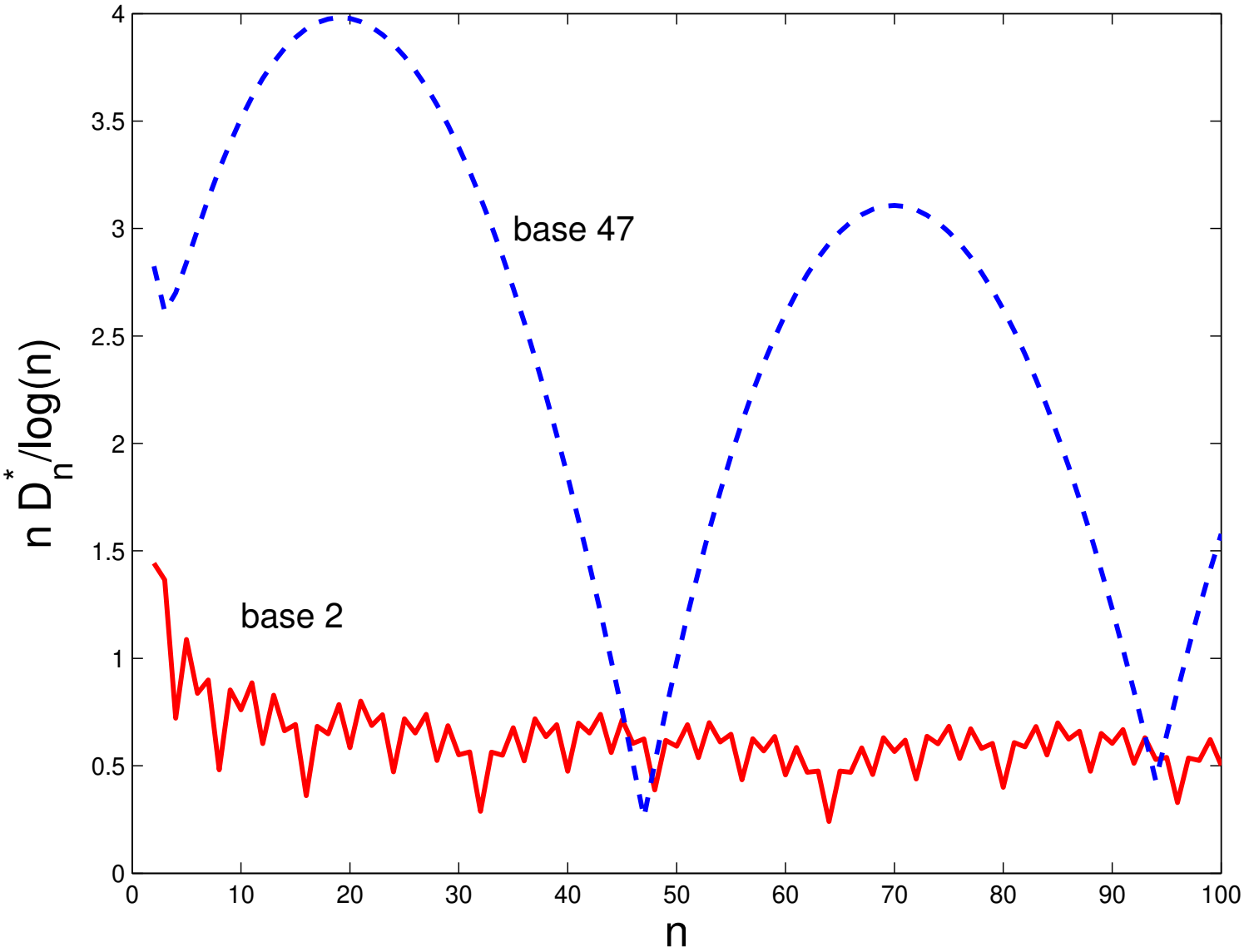
Discrepancy:

$$D_n^*(\mathbf{X}_\infty) \leq C_d \frac{(\log n)^d}{n} + \mathcal{O}\left(\frac{(\log n)^{d-1}}{n}\right)$$

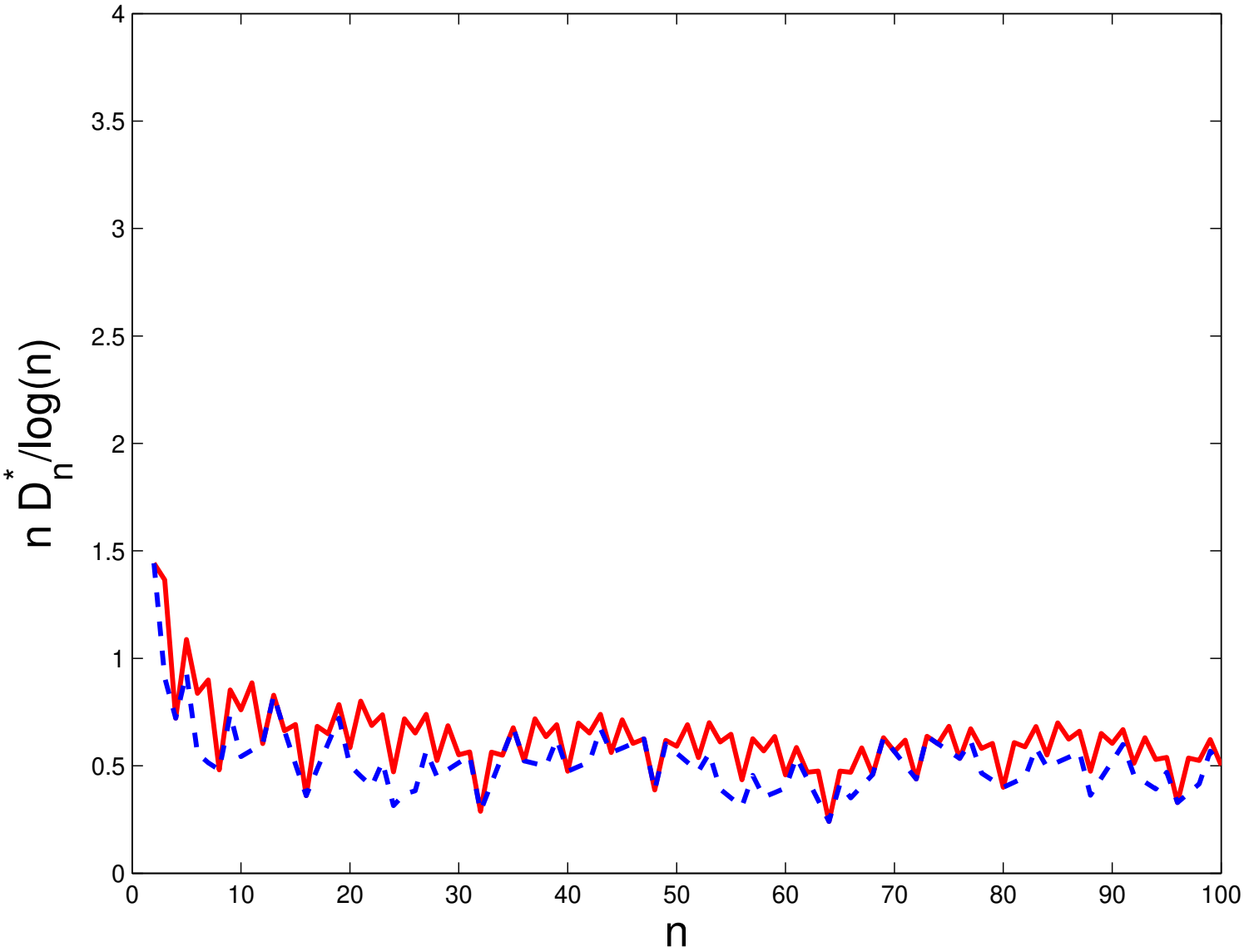
bound similar to that of Halton LDS

. . . but here $C_d \searrow 0$ super-exponentially fast as $d \to \infty$!

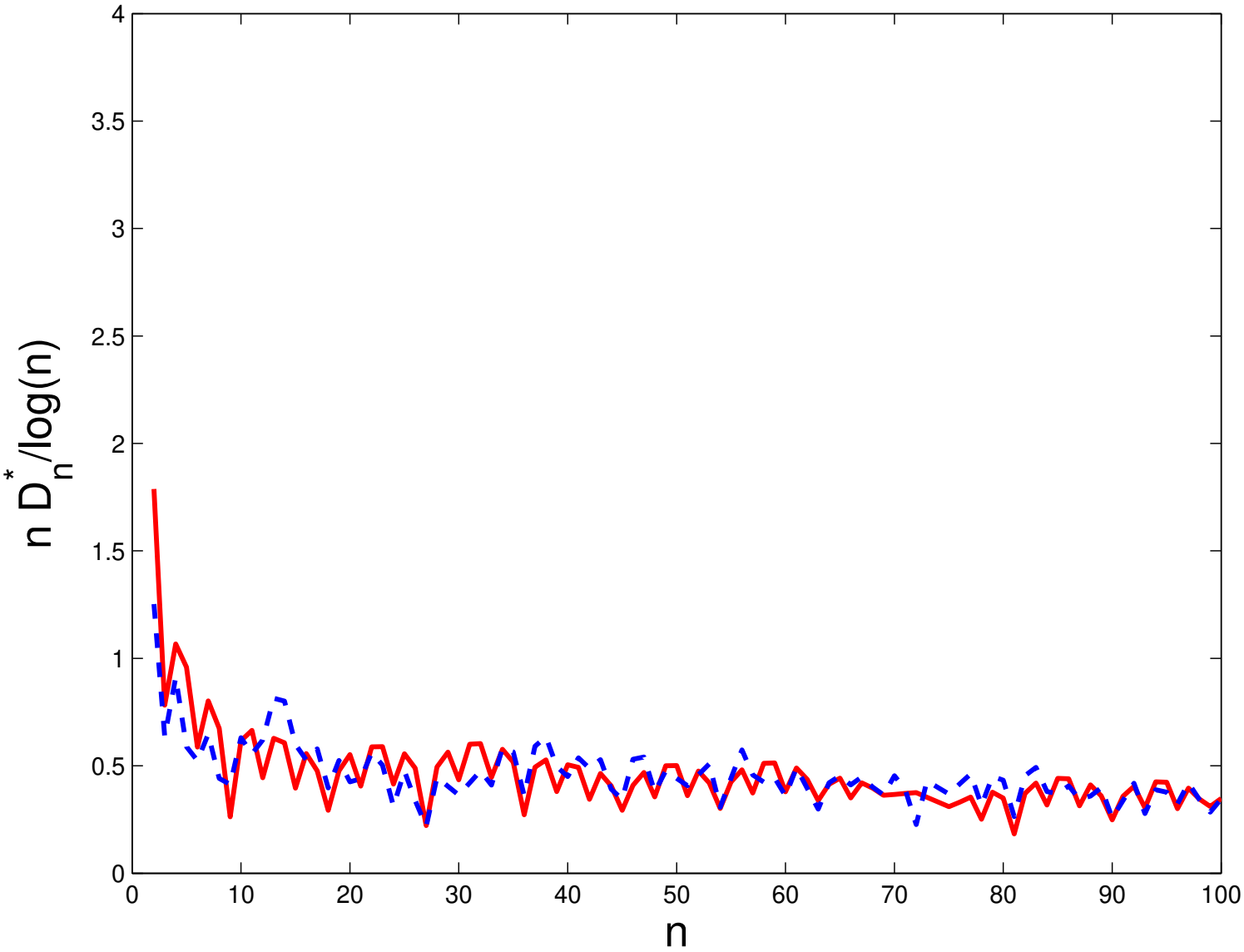
Halton: $(n/\log n) \times D_n^*$ for $\{\mathbf{X}_n\}_1$ and $\{\mathbf{X}_n\}_{15}$



Sobol' : $(n/\log n) \times D_n^*$ for $\{\mathbf{X}_n\}_1$ and $\{\mathbf{X}_n\}_{15}$

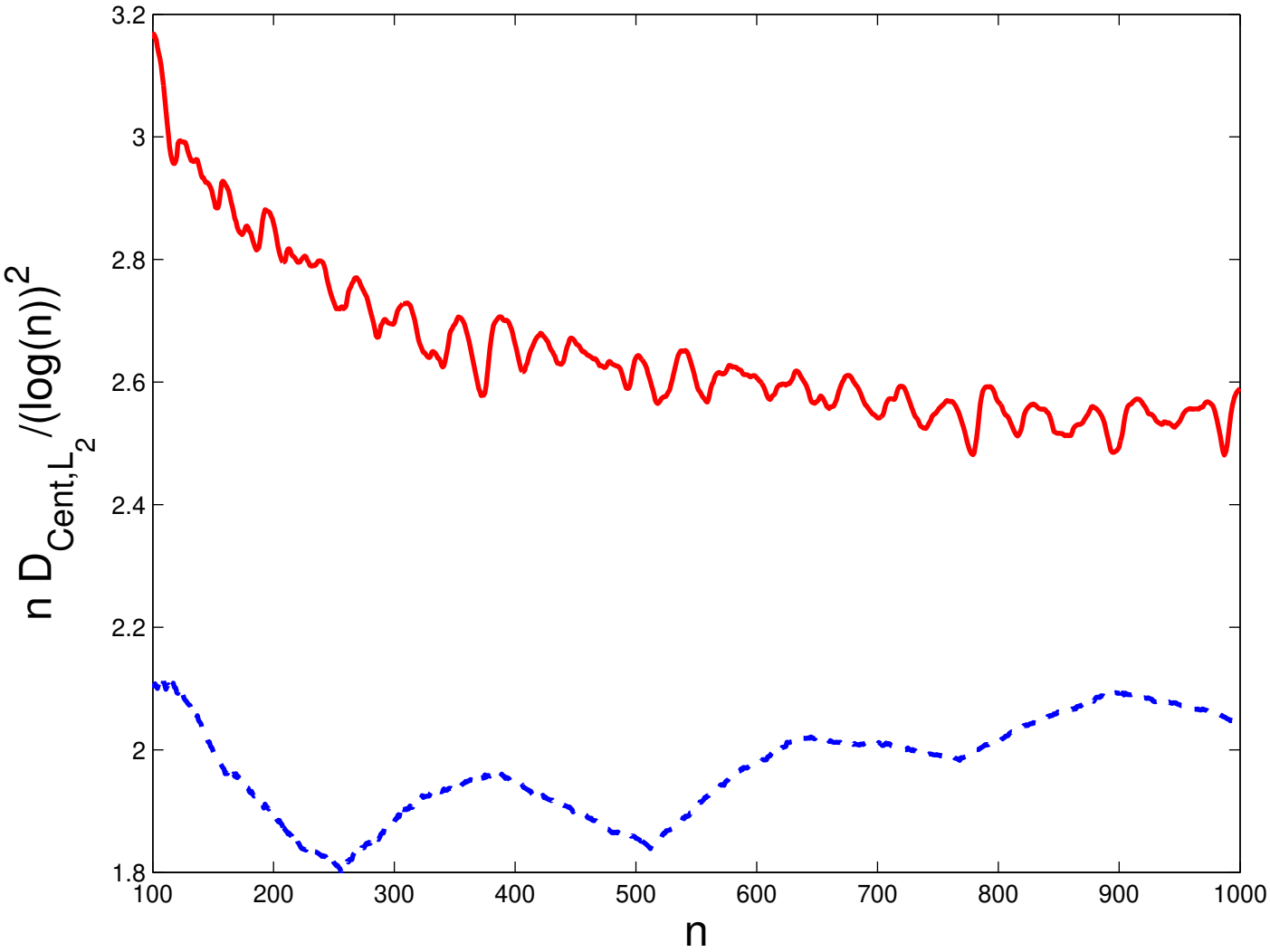


Faure : $(n/\log n) \times D_n^*$ for $\{\mathbf{X}_n\}_1$ and $\{\mathbf{X}_n\}_{15}$



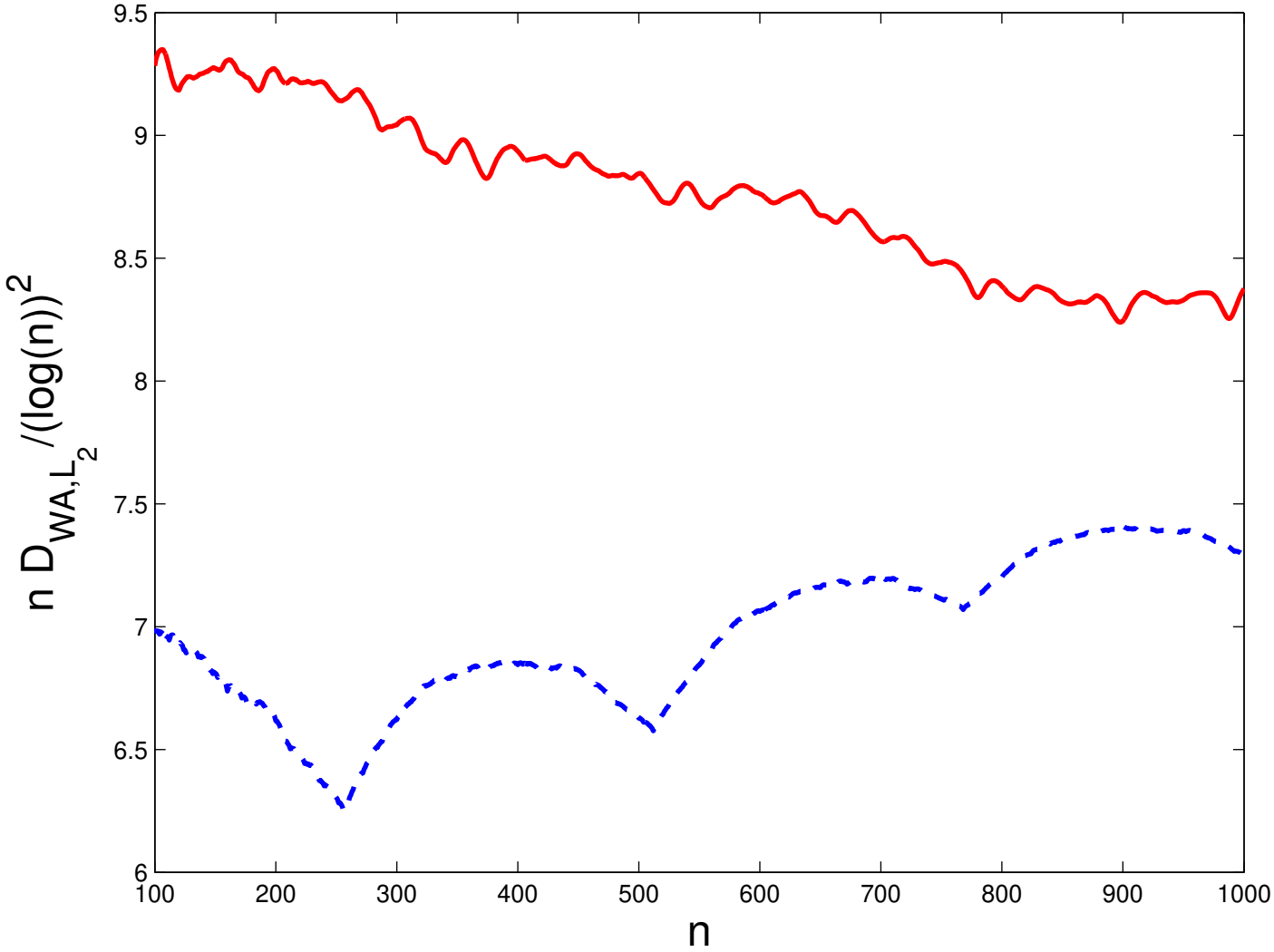
$[n/(\log n)^2] \times D_{Cent,L_2}(\mathbf{X}_n)$ for Halton and Sobol' $(d = 15)$

$(/(\log n)^2$, but no obvious normalization for $D_{Cent,L_2})$



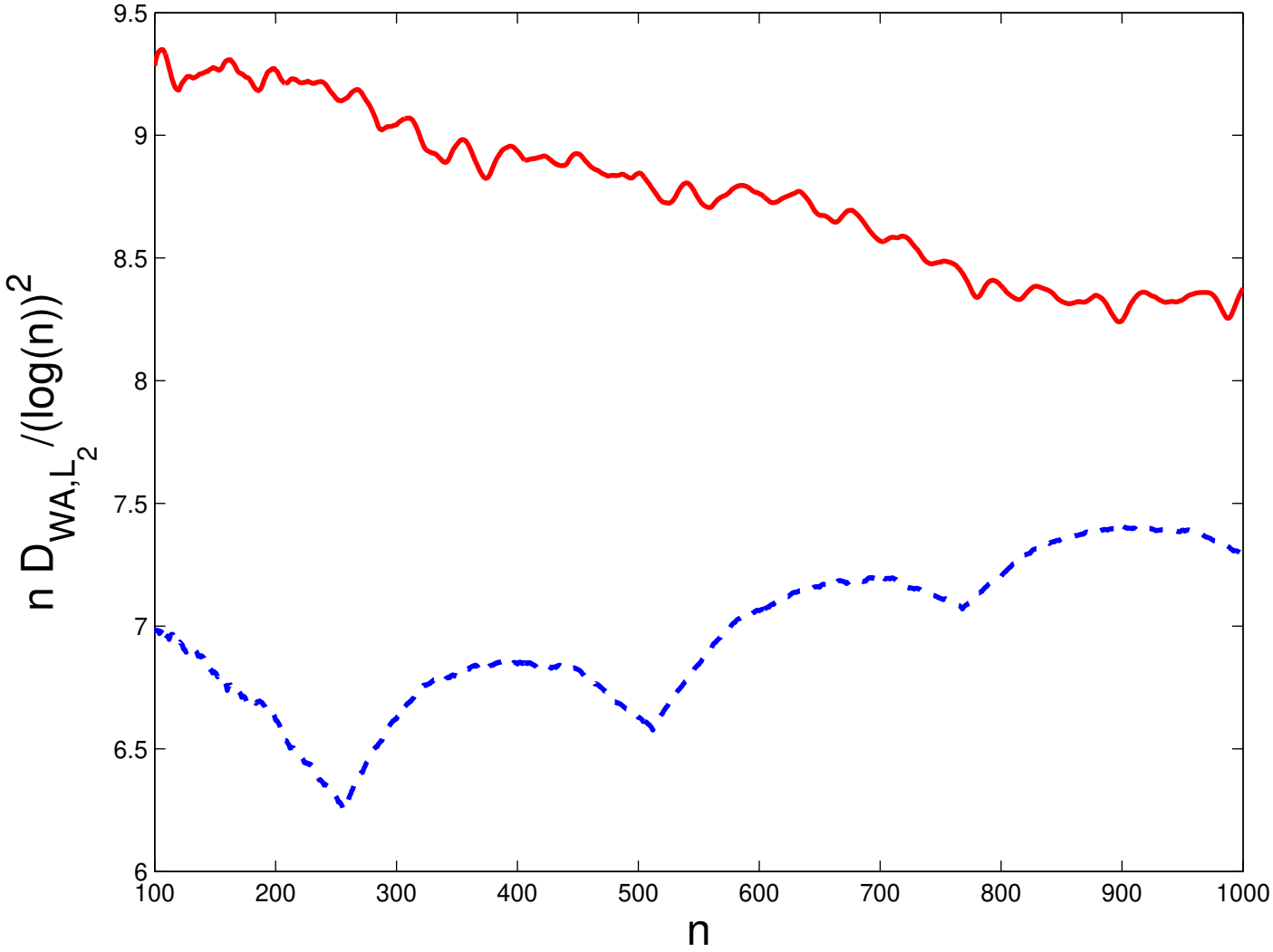
$[n/(\log n)^2] \times D_{WA,L_2}(\mathbf{X}_n)$ for Halton and Sobol' $(d = 15)$

$(/(\log n)^2$, but no obvious normalization for $D_{WA,L_2})$



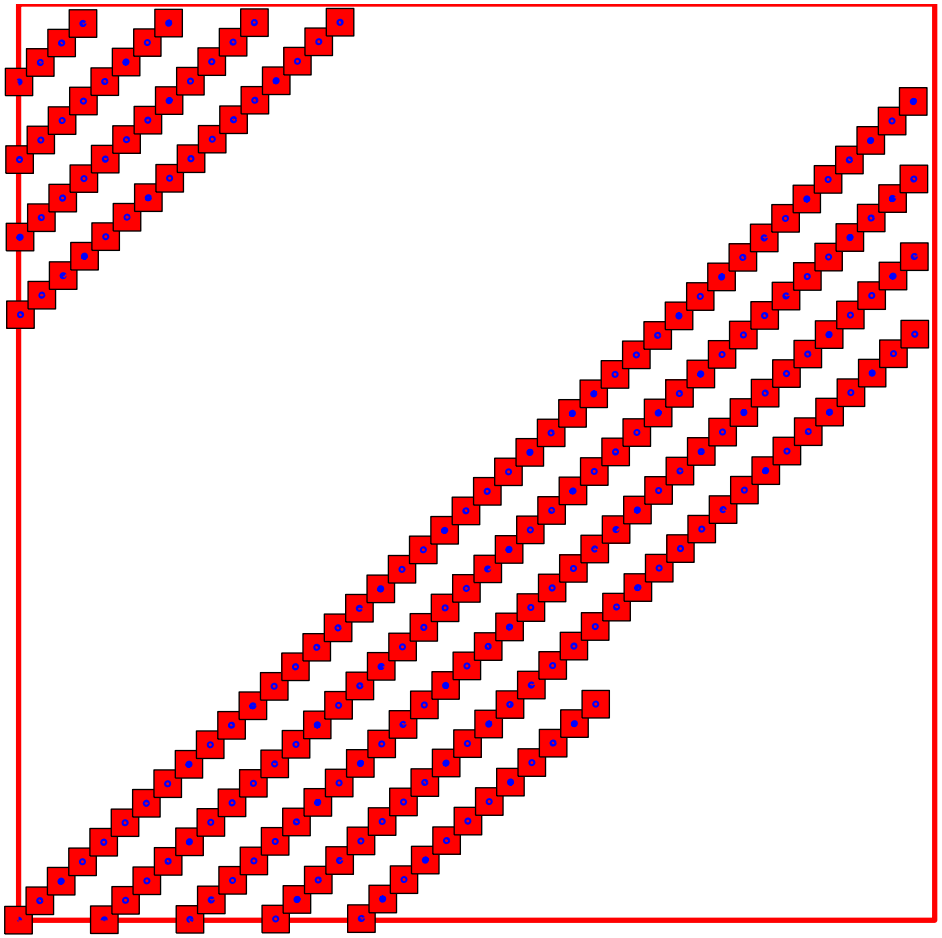
$[n/(\log n)^2] \times D_{WA,L_2}(\mathbf{X}_n)$ for Halton and Sobol' $(d = 15)$

$(/(\log n)^2$, but no obvious normalization for $D_{WA,L_2})$

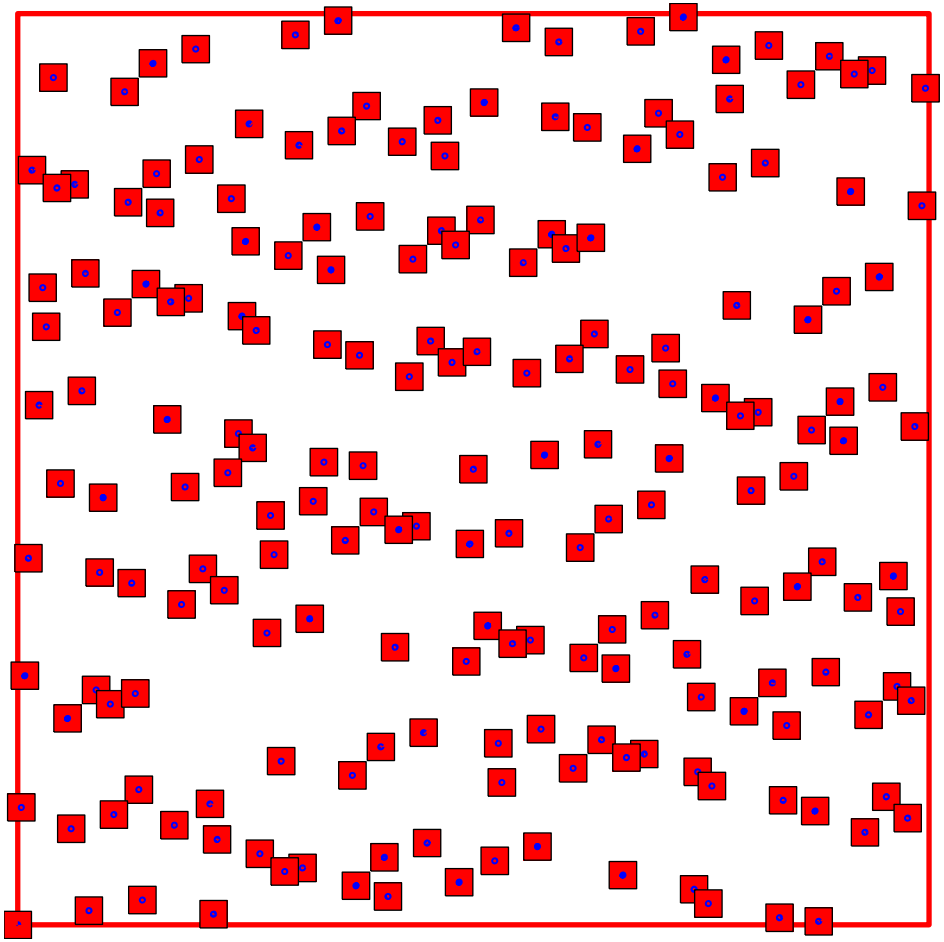


➤ $n = 1\,000$ is small: the cube $[0, 1]^{15}$ has $2^{15} = 32\,768$ vertices!

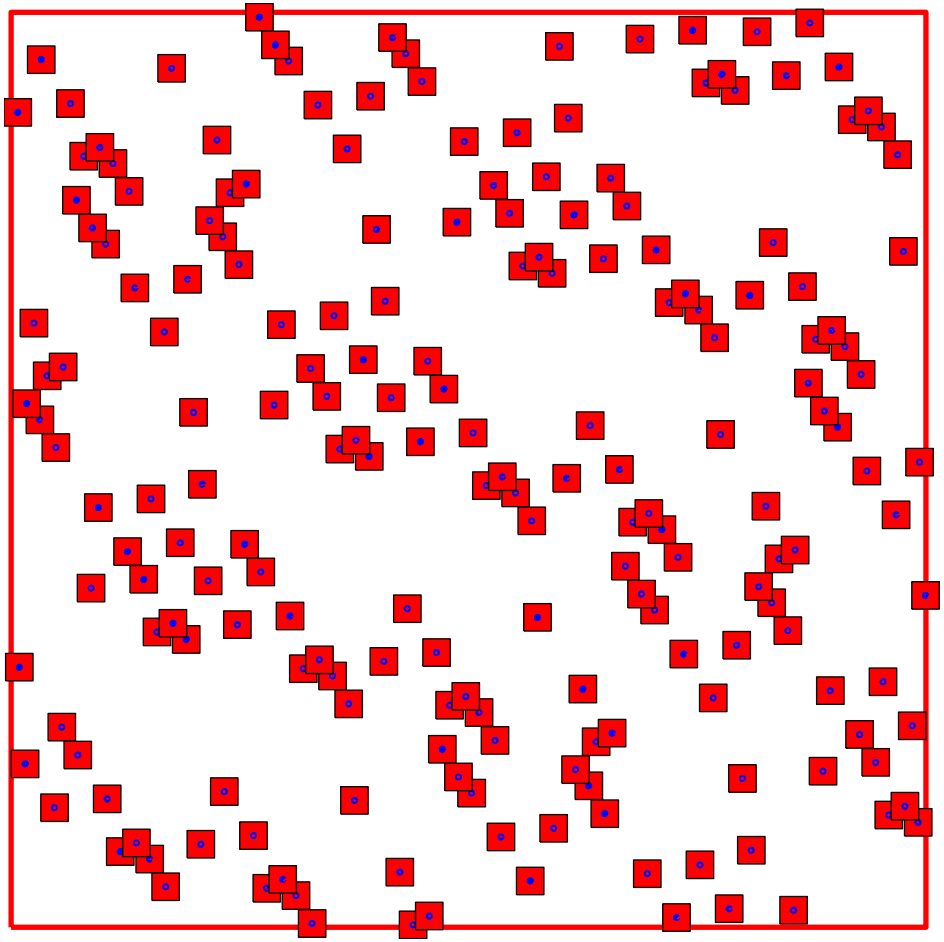
Halton, $d = 15$: $\{\mathbf{x}_k\}_{14}$ and $\{\mathbf{x}_k\}_{15}$, $k = 1, \ldots, 200$



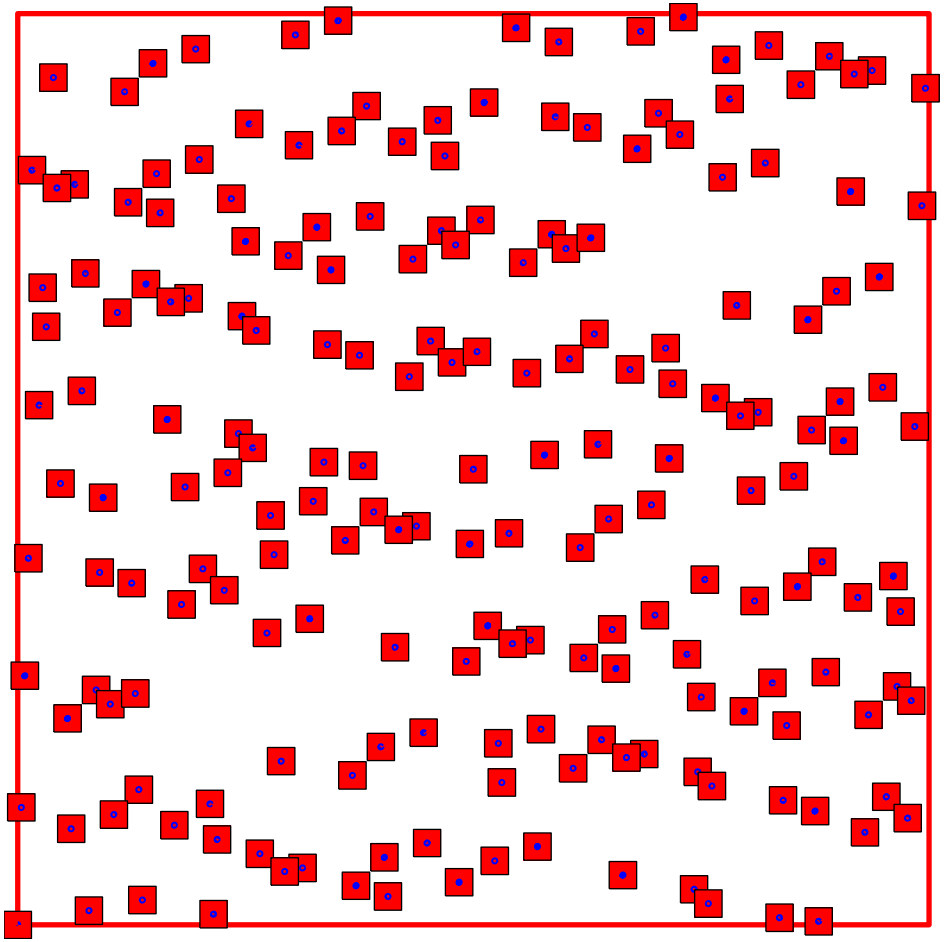
Sobol': $d = 15$, $\{\mathbf{x}_k\}_{14}$ and $\{\mathbf{x}_k\}_{15}$, $k = 1, \ldots, 200$



Faure, $d = 15$: $\{\mathbf{x}_k\}_{14}$ and $\{\mathbf{x}_k\}_{15}$, $k = 1, \ldots, 200$



Sobol': $d = 15$, $\{\mathbf{x}_k\}_{14}$ and $\{\mathbf{x}_k\}_{15}$, $k = 1, \ldots, 200$



3 Dispersion & miniMax

3.1 **Dispersion** (Niederreiter, 1992, Chap. 6)

Discrepancy measures uniformity of the distribution of the $\mathbf{x}_k$

➠ we can also restrict our attention to the "filling" of $\mathscr{X}$ by $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$

$$\boxed{d_n(\mathbf{X}_n, \mathscr{X}) \triangleq \sup_{\mathbf{x} \in \mathscr{X}} \min_{1 \leq k \leq n} \Delta(\mathbf{x}, \mathbf{x}_k)}$$

(▲ we shall minimize this measure of dispersion ▲)

3 Dispersion & miniMax

3.1 **Dispersion** (Niederreiter, 1992, Chap. 6)

Discrepancy measures uniformity of the distribution of the $\mathbf{x}_k$

➠ we can also restrict our attention to the "filling" of $\mathscr{X}$ by $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$

$$\boxed{d_n(\mathbf{X}_n, \mathscr{X}) \triangleq \sup_{\mathbf{x} \in \mathscr{X}} \min_{1 \leq k \leq n} \Delta(\mathbf{x}, \mathbf{x}_k)}$$

(▲ we shall minimize this measure of dispersion ▲)

- If $\Delta(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|$ (Euclidean distance),

  ➠ $\boxed{d_n(\mathbf{X}_n, \mathscr{X}) = \phi_{mM}(\mathbf{X}_n) = \text{miniMax distance criterion}}$

3 Dispersion & miniMax

3.1 **Dispersion** (Niederreiter, 1992, Chap. 6)

Discrepancy measures uniformity of the distribution of the $\mathbf{x}_k$

➠ we can also restrict our attention to the "filling" of $\mathscr{X}$ by $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$

$$\boxed{d_n(\mathbf{X}_n, \mathscr{X}) \triangleq \sup_{\mathbf{x} \in \mathscr{X}} \min_{1 \leq k \leq n} \Delta(\mathbf{x}, \mathbf{x}_k)}$$

(▲ we shall minimize this measure of dispersion ▲)

- If $\Delta(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|$ (Euclidean distance),
  ➠ $\boxed{d_n(\mathbf{X}_n, \mathscr{X}) = \phi_{mM}(\mathbf{X}_n) = \text{miniMax distance criterion}}$
- If $\Delta(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|_\infty = \max_{1 \leq j \leq d} |\{\mathbf{x}\}_j - \{\mathbf{x}'\}_j|$ (ℓ_∞ distance),
  ➠ $d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$ ("balls" are cubes, easier to pack)

3 Dispersion & miniMax

3.1 **Dispersion** (Niederreiter, 1992, Chap. 6)

Discrepancy measures uniformity of the distribution of the $\mathbf{x}_k$

➠ we can also restrict our attention to the "filling" of $\mathscr{X}$ by $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$

$$\boxed{d_n(\mathbf{X}_n, \mathscr{X}) \triangleq \sup_{\mathbf{x} \in \mathscr{X}} \min_{1 \leq k \leq n} \Delta(\mathbf{x}, \mathbf{x}_k)}$$

(▲ we shall minimize this measure of dispersion ▲)

- If $\Delta(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|$ (Euclidean distance),
  ➠ $\boxed{d_n(\mathbf{X}_n, \mathscr{X}) = \phi_{mM}(\mathbf{X}_n) = \text{miniMax distance criterion}}$
- If $\Delta(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|_\infty = \max_{1 \leq j \leq d} |\{\mathbf{x}\}_j - \{\mathbf{x}'\}_j|$ (ℓ_∞ distance),
  ➠ $d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$ ("balls" are cubes, easier to pack)
- $d_{\infty,n}(\mathbf{X}_n, \mathscr{X}) = d_n(\mathbf{X}_n, \mathscr{X})$ for $d = 1$

- $d_{\infty,n}(\mathbf{X}_n, \mathscr{X}) \leq d_n(\mathbf{X}_n, \mathscr{X}) \leq \sqrt{d}\, d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$

- $d_{\infty,n}(\mathbf{X}_n, \mathscr{X}) \leq d_n(\mathbf{X}_n, \mathscr{X}) \leq \sqrt{d}\, d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$
- Sphere covering (§ 1.1 and 1.5) :
  - $\left(\frac{\text{vol}(\mathscr{X})}{V_d}\right)^{1/d} \frac{1}{n^{1/d}} \leq d_n(\mathbf{X}_n, \mathscr{X}) = \Phi_{mM}(\mathbf{X}_n)$ $(V_d = \text{vol}[\mathscr{B}(0,1)])$

- $d_{\infty,n}(\mathbf{X}_n, \mathscr{X}) \leq d_n(\mathbf{X}_n, \mathscr{X}) \leq \sqrt{d}\, d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$
- Sphere covering (§ 1.1 and 1.5) :
  - $\left(\frac{\text{vol}(\mathscr{X})}{V_d}\right)^{1/d} \frac{1}{n^{1/d}} \leq d_n(\mathbf{X}_n, \mathscr{X}) = \Phi_{mM}(\mathbf{X}_n)$ $(V_d = \text{vol}[\mathscr{B}(0,1)])$
- Cube covering : $n[2\, d_{\infty,n}(\mathbf{X}_n, \mathscr{X})]^d \geq \text{vol}(\mathscr{X})$
  - $\frac{1}{2}\, (\text{vol}(\mathscr{X}))^{1/d}\, \frac{1}{n^{1/d}} \leq d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$

- $d_{\infty,n}(\mathbf{X}_n, \mathscr{X}) \leq d_n(\mathbf{X}_n, \mathscr{X}) \leq \sqrt{d}\, d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$
- Sphere covering (§ 1.1 and 1.5) :
  - ➠ $\left(\frac{\text{vol}(\mathscr{X})}{V_d}\right)^{1/d} \frac{1}{n^{1/d}} \leq d_n(\mathbf{X}_n, \mathscr{X}) = \Phi_{mM}(\mathbf{X}_n)$ $(V_d = \text{vol}[\mathscr{B}(0,1)])$
- Cube covering : $n[2\, d_{\infty,n}(\mathbf{X}_n, \mathscr{X})]^d \geq \text{vol}(\mathscr{X})$
  - ➠ $\frac{1}{2}\, (\text{vol}(\mathscr{X}))^{1/d}\, \frac{1}{n^{1/d}} \leq d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$

For $\mathscr{X} = [0,1]^d$ $(\text{vol}(\mathscr{X}) = 1)$

$$\frac{1}{2}\, \frac{1}{\lfloor n^{1/d} \rfloor} \leq d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$$

with equality for some $\mathbf{X}_n$, for any n an d

- $d_{\infty,n}(\mathbf{X}_n, \mathscr{X}) \leq d_n(\mathbf{X}_n, \mathscr{X}) \leq \sqrt{d}\, d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$
- Sphere covering (§ 1.1 and 1.5) :
  - ➠ $\left(\frac{\text{vol}(\mathscr{X})}{V_d}\right)^{1/d} \frac{1}{n^{1/d}} \leq d_n(\mathbf{X}_n, \mathscr{X}) = \Phi_{mM}(\mathbf{X}_n)$ $(V_d = \text{vol}[\mathscr{B}(0,1)])$
- Cube covering : $n[2\, d_{\infty,n}(\mathbf{X}_n, \mathscr{X})]^d \geq \text{vol}(\mathscr{X})$
  - ➠ $\frac{1}{2}\, (\text{vol}(\mathscr{X}))^{1/d}\, \frac{1}{n^{1/d}} \leq d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$

For $\mathscr{X} = [0,1]^d$ $(\text{vol}(\mathscr{X}) = 1)$

$$\frac{1}{2}\, \frac{1}{\lfloor n^{1/d} \rfloor} \leq d_{\infty,n}(\mathbf{X}_n, \mathscr{X}) \leq [D_n(\mathbf{X}_n)]^{1/d} \leq 2[D_n^*(\mathbf{X}_n)]^{1/d}$$

with equality for some $\mathbf{X}_n$, for any n an d

$$\leq A \frac{(\log n)^{(d-1)/d}}{n^{1/d}} \text{ for } \mathbf{X}_n \text{ a LDS}$$

⏩ low discrepancy ⇒ low dispersion

- $d_{\infty,n}(\mathbf{X}_n, \mathscr{X}) \leq d_n(\mathbf{X}_n, \mathscr{X}) \leq \sqrt{d}\, d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$
- Sphere covering (§ 1.1 and 1.5) :
  - ➠ $\left(\frac{\text{vol}(\mathscr{X})}{V_d}\right)^{1/d} \frac{1}{n^{1/d}} \leq d_n(\mathbf{X}_n, \mathscr{X}) = \Phi_{mM}(\mathbf{X}_n)$ $(V_d = \text{vol}[\mathscr{B}(0,1)])$
- Cube covering : $n[2\, d_{\infty,n}(\mathbf{X}_n, \mathscr{X})]^d \geq \text{vol}(\mathscr{X})$
  - ➠ $\frac{1}{2}\, (\text{vol}(\mathscr{X}))^{1/d}\, \frac{1}{n^{1/d}} \leq d_{\infty,n}(\mathbf{X}_n, \mathscr{X})$

For $\mathscr{X} = [0,1]^d$ $(\text{vol}(\mathscr{X}) = 1)$

$$\frac{1}{2}\, \frac{1}{\lfloor n^{1/d} \rfloor} \leq d_{\infty,n}(\mathbf{X}_n, \mathscr{X}) \leq [D_n(\mathbf{X}_n)]^{1/d} \leq 2[D_n^*(\mathbf{X}_n)]^{1/d}$$

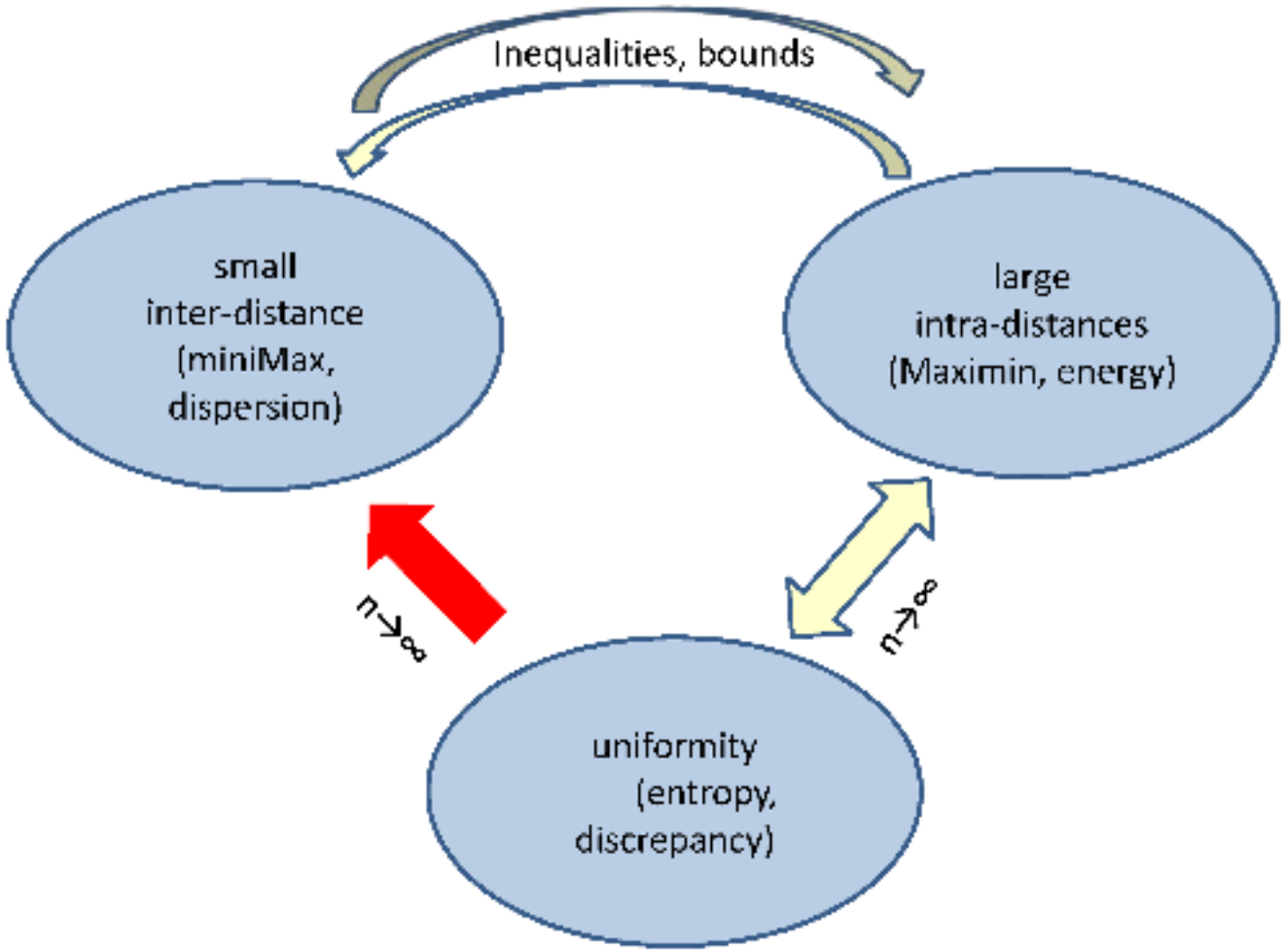
with equality for some $\mathbf{X}_n$, for any n an d

$$\leq A \frac{(\log n)^{(d-1)/d}}{n^{1/d}} \text{ for } \mathbf{X}_n \text{ a LDS}$$

⏩ low discrepancy ⇒ low dispersion

$$\Rightarrow \max\left\{ \frac{1}{(nV_d)^{1/d}}, \frac{1}{2}\, \frac{1}{\lfloor n^{1/d} \rfloor} \right\} \leq \underbrace{= \Phi^*_{mM,n}}_{\min_{\mathbf{X}_n} \Phi_{mM}(\mathbf{X}_n)} \underbrace{\leq \frac{\sqrt{d}}{2}\, \frac{1}{\lfloor n^{1/d} \rfloor}}_{\S 1.5}$$

(slightly improves the bounds of §1.5)


Inequalities, bounds
small
inter-distance
(miniMax,
dispersion)
large
intra-distances
(Maximin, energy)
$n \to \infty$
$n \to \infty$
uniformity
(entropy,
discrepancy)

3.2 **Low dispersion sequences** (Niederreiter, 1992, Chap. 6)

$\underline{d = 1}$: for any sequence $\mathbf{X}_\infty$, $\limsup_{n\to\infty} n d_n(\mathbf{X}_\infty) \geq \frac{1}{2\log 2} \simeq 0.7213$

3.2 **Low dispersion sequences** (Niederreiter, 1992, Chap. 6)

$\underline{d = 1}$: for any sequence $\mathbf{X}_\infty$, $\limsup_{n\to\infty} n d_n(\mathbf{X}_\infty) \geq \frac{1}{2\log 2} \simeq 0.7213$

The bound is reached for Ruzsa sequence:

$$x_1 = 1,\ x_k = \left\{\frac{\log(2k-3)}{\log 2}\right\},\ k \geq 2$$

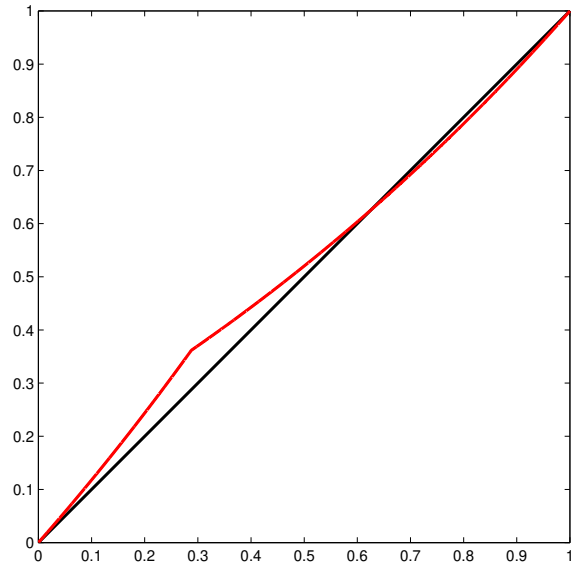
3.2 **Low dispersion sequences** (Niederreiter, 1992, Chap. 6)

$d = 1$: for any sequence $\mathbf{X}_\infty$, $\limsup_{n\to\infty} n d_n(\mathbf{X}_\infty) \geq \frac{1}{2\log 2} \simeq 0.7213$

The bound is reached for Ruzsa sequence:

$$x_1 = 1,\ x_k = \left\{\frac{\log(2k-3)}{\log 2}\right\},\ k \geq 2$$

but $D_n^*(\mathbf{X}_\infty) \not\to 0$! F_n for $n = 10\,000$ points ($D_n^* \simeq 0.0740$)



Low discrepancy ($\Leftrightarrow$ uniformity) $\Rightarrow$ low dispersion

but low dispersion $\not\Rightarrow$ low discrepancy

$\underline{d > 1}$: $\frac{1}{2} \leq \boxed{\inf_{\mathbf{X}_\infty} \limsup_{n\to\infty} n^{1/d} d_{\infty,n}(\mathbf{X}_\infty)} \leq \frac{1}{2\log 2}$

Sequences that reach the upper bound $\frac{1}{2\log 2}$ are known

The smallest value $\inf_{\mathbf{X}_\infty} \limsup_{n\to\infty} n^{1/d} d_{\infty,n}(\mathbf{X}_\infty)$ is unknown

(and best sequences $\mathbf{X}_\infty$ are unknown too)

(very) little is known about $n^{1/d} d_n(\mathbf{X}_\infty) = n^{1/d} \Phi_{mM}(\mathbf{X}_\infty)$!

$\underline{d > 1}$: $\frac{1}{2} \leq \boxed{\inf_{\mathbf{X}_\infty} \limsup_{n\to\infty} n^{1/d} d_{\infty,n}(\mathbf{X}_\infty)} \leq \frac{1}{2\log 2}$

Sequences that reach the upper bound $\frac{1}{2\log 2}$ are known

The smallest value $\inf_{\mathbf{X}_\infty} \limsup_{n\to\infty} n^{1/d} d_{\infty,n}(\mathbf{X}_\infty)$ is unknown

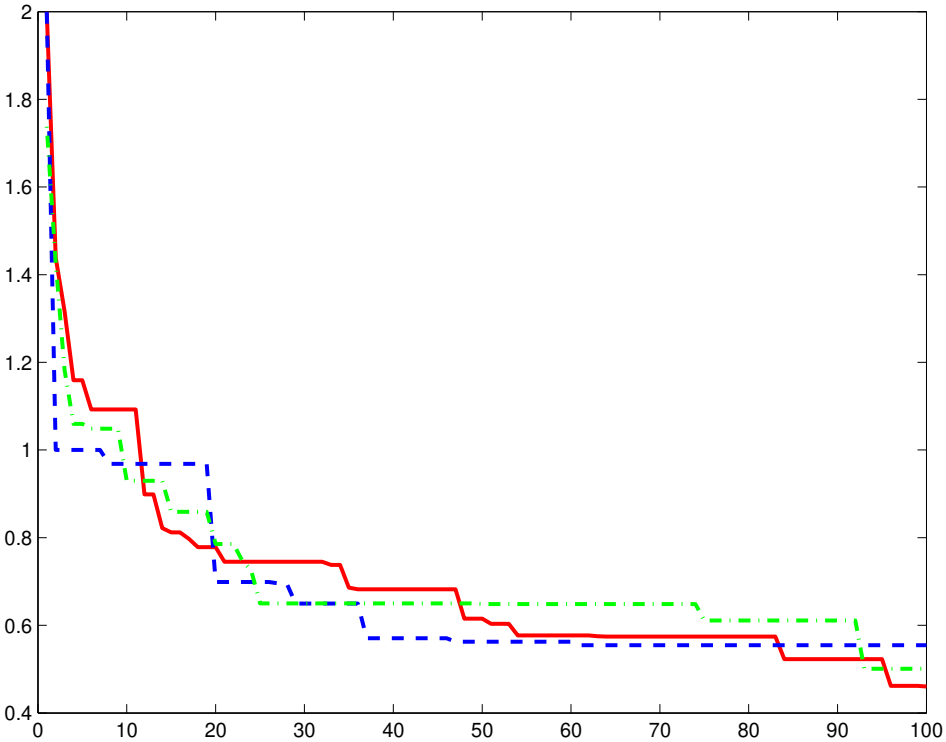
(and best sequences $\mathbf{X}_\infty$ are unknown too)

(very) little is known about $n^{1/d} d_n(\mathbf{X}_\infty) = n^{1/d} \Phi_{mM}(\mathbf{X}_\infty)$!

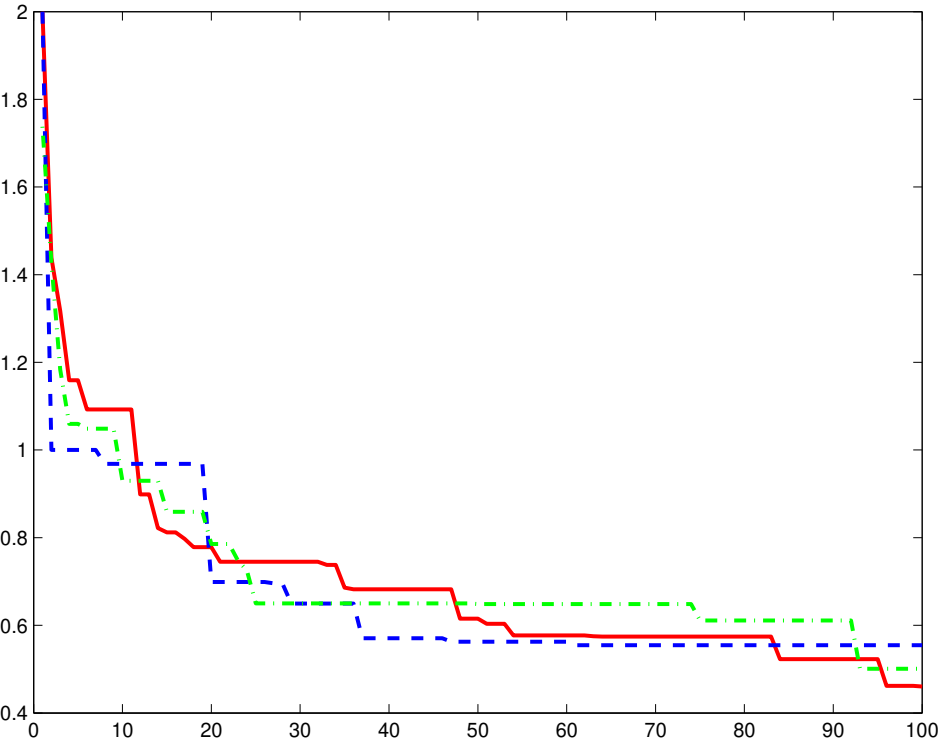
Upper bounds (rather pessimistic):

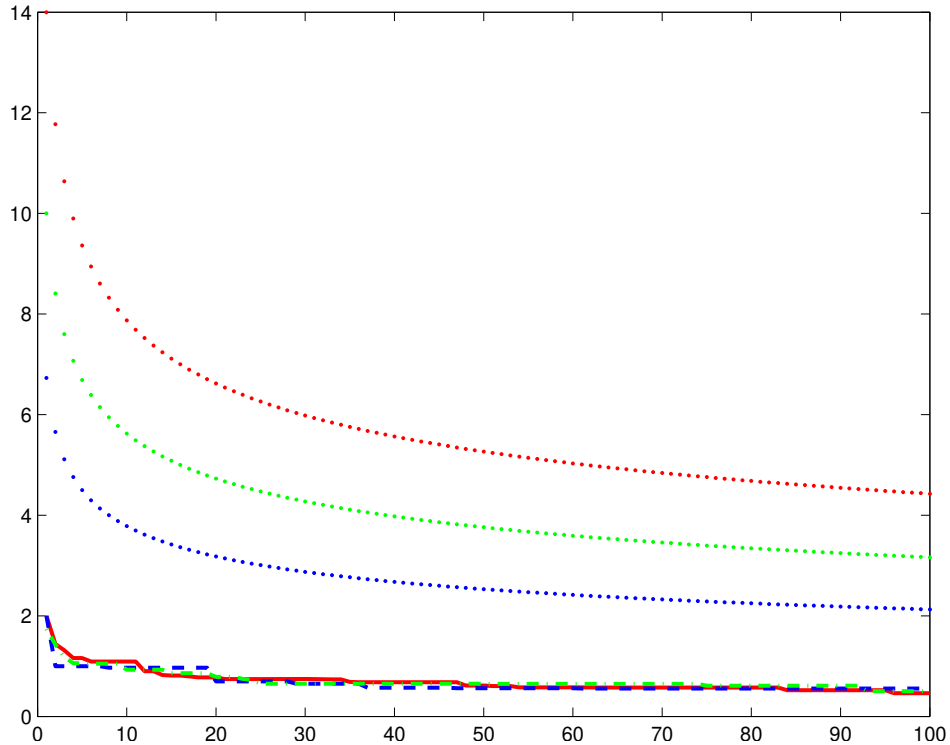
- Halton in base $(b_1, \ldots, b_d)$ ⇛ $\boxed{d_{\infty,n}(\mathbf{X}_\infty) < \frac{\max_{1\leq i\leq d} b_i}{n^{1/d}}}$
- (t,d)-sequence in base b ⇛ $\boxed{d_{\infty,n}(\mathbf{X}_\infty) < \frac{b^{1+t/d}}{n^{1/d}}}$

Ex. : $d = 4$, n from 1 to 100, $\Phi_{mM}(\mathbf{X}_n)$ for Halton ($\mathbf{b} = (2, 3, 5, 7)$), Sobol ($t = 3$, $b = 2$) and Faure ($t = 0$, $b = 5$)

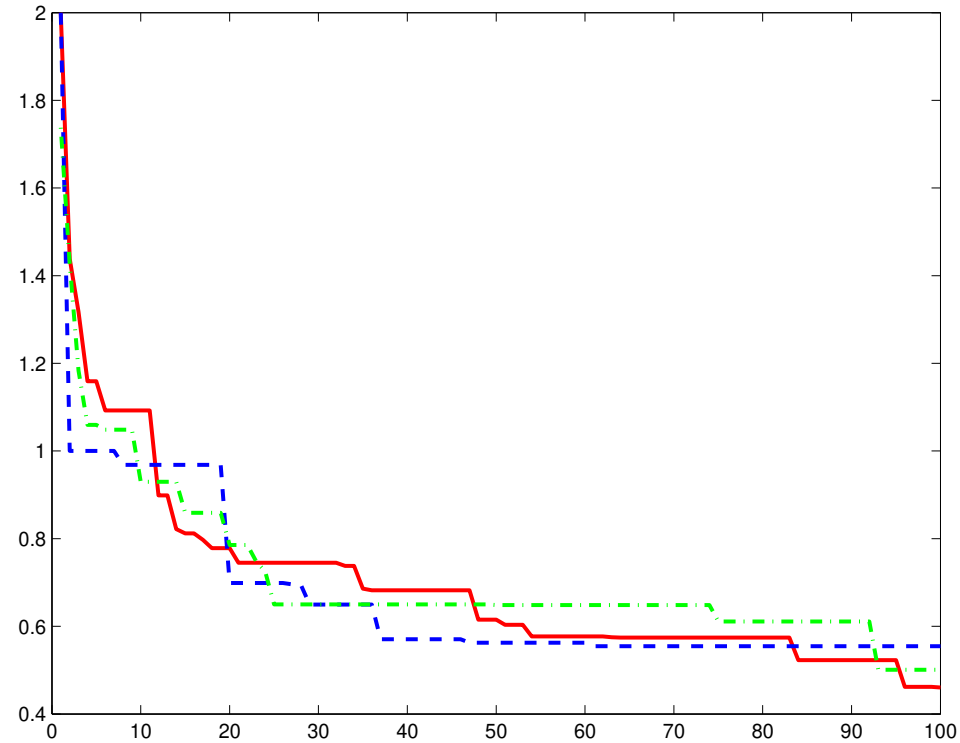


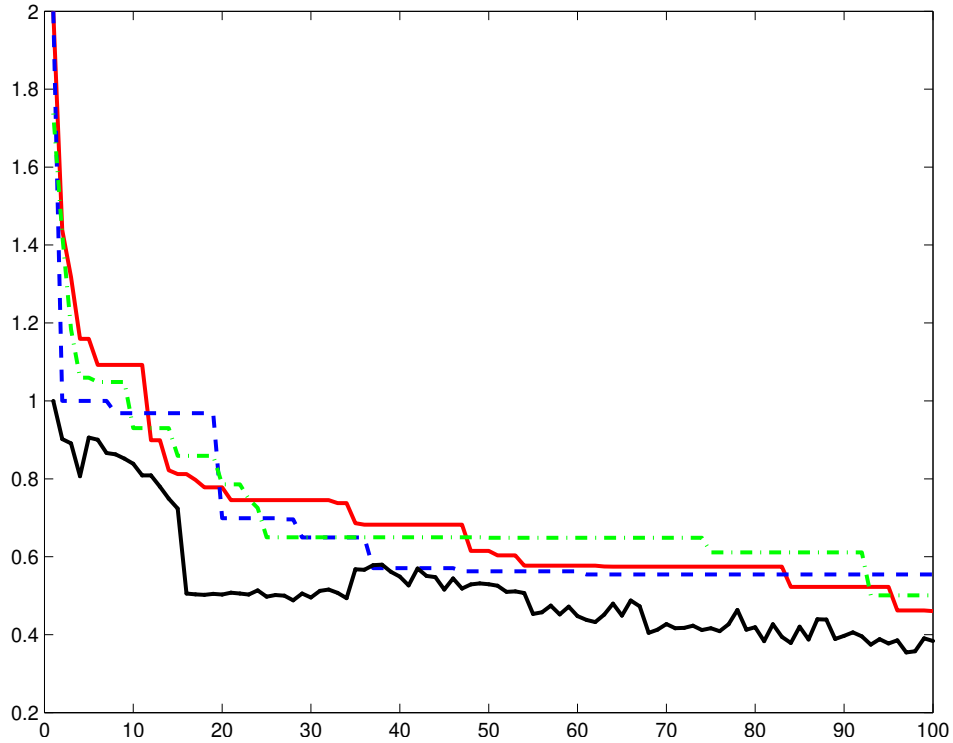
Ex. : $d = 4$, n from 1 to 100, $\Phi_{mM}(\mathbf{X}_n)$ for Halton ($\mathbf{b} = (2, 3, 5, 7)$), Sobol ($t = 3$, $b = 2$) and Faure ($t = 0$, $b = 5$)





Ex. : $d = 4$, n from 1 to 100, $\Phi_{mM}(\mathbf{X}_n)$ for Halton ($\mathbf{b} = (2, 3, 5, 7)$), Sobol ($t = 3$, $b = 2$) and Faure ($t = 0$, $b = 5$)





running `kmeans` algorithm on 10 000 points of Sobol' LDS (with centers initialized at the first n points)

4 Conclusions part (1) — without model

- Many design criteria available (geometry, uniformity)

4 Conclusions part (1) — without model

- Many design criteria available (geometry, uniformity)
- Their optimization is difficult (non convex, multimodal, sometimes non differentiable) ➠ workable d not too large

  $\Phi_{mM}(\cdot)$ is rather compelling, its evaluation is not trivial but possible

  can be optimized by clustering (with Chebyshev centers) for small d,

  by stochastic approximation otherwise

4 Conclusions part (1) — without model

- Many design criteria available (geometry, uniformity)
- Their optimization is difficult (non convex, multimodal, sometimes non differentiable) ➠ workable d not too large

  $\Phi_{mM}(\cdot)$ is rather compelling, its evaluation is not trivial but possible

  can be optimized by clustering (with Chebyshev centers) for small d,

  by stochastic approximation otherwise
- Low discrepancy sequences:

  easy to generate

  the sequence is well distributed (not necessary to choose n *a priori*)

  can be used for any compact $\mathscr{X}$ (with non empty interior)

  (generate points in a cube containing $\mathscr{X}$, and reject points not in $\mathscr{X}$)

The curse of dimensionality is always present!

$d = 50$, Faure $(0, d)$-sequence ➠ b prime $\geq d \rightarrow b = 53$

If we want to ensure that there is a point in each box cut along q dimensions, then

$\sum_{j=1}^{d} q_j = q = m - t = m$

➠ $n = b^q$

$q = 2$ ➠ $n = 2809$

$q = 50$ ➠ $n \simeq 1.6360\ 10^{86}$

The curse of dimensionality is always present!

$d = 50$, Faure $(0, d)$-sequence ➠ b prime $\geq d \rightarrow b = 53$
If we want to ensure that there is a point in each box cut along q dimensions, then $\sum_{j=1}^{d} q_j = q = m - t = m$
➠ $n = b^q$

$q = 2$ ➠ $n = 2809$
$q = 50$ ➠ $n \simeq 1.6360\ 10^{86}$

$d = 50$, (t, d)-sequence in base 2, with smallest possible $t \rightarrow t = 77$
If we want that each elementary interval cut along each dimension contains some points, then $n = b^m$ with $\sum_{j=1}^{d} q_j = d = m - t$ ➠ $m = 127$ and $n \simeq 1.7014\ 10^{38}$

The curse of dimensionality is always present!

$d = 50$, Faure $(0, d)$-sequence ➠ b prime $\geq d \rightarrow b = 53$
If we want to ensure that there is a point in each box cut along q dimensions, then $\sum_{j=1}^{d} q_j = q = m - t = m$
➠ $n = b^q$

$q = 2$ ➠ $n = 2809$
$q = 50$ ➠ $n \simeq 1.6360\ 10^{86}$

$d = 50$, (t, d)-sequence in base 2, with smallest possible $t \rightarrow t = 77$
If we want that each elementary interval cut along each dimension contains some points, then $n = b^m$ with $\sum_{j=1}^{d} q_j = d = m - t$ ➠ $m = 127$ and $n \simeq 1.7014\ 10^{38}$

See (Owen, 1998) for possible constructions, such as

$$\mathbf{X}_n = \{\mathbf{X}_n\}_{1:d} = \begin{pmatrix} \{\mathbf{X}_n\}_{1:s} \\ \{\mathbf{X}_n\}_{s+1:d} \end{pmatrix}$$

with e.g. $\{\mathbf{X}_n\}_{1:s}$ a LDS sequence
$\{\mathbf{X}_n\}_{s+1:d}$ a Lh

References I

Audze, P., Eglais, V., 1977. New approach for planning out of experiments. Problems of Dynamics and Strengths 35, 104–107.

Bates, R., Buck, R., Riccomagno, E., Wynn, H., 1996. Experimental design and observation for large systems. Journal of the Royal Statistical Society. Series B (Methodological), 77–94.

Beardwood, J., Halton, J., Hammersley, J., 1959. The shortest path through many points. Mathematical Proceedings of the Cambridge Philosophical Society 55 (4), 299–327.

Biedermann, S., Dette, H., June 2001. Minimax optimal designs for nonparametric regression — a further optimality property of the uniform distribution. In: A.C. Atkinson, P. H., Müller, W. (Eds.), mODa'6 – Advances in Model–Oriented Design and Analysis, Proceedings of the 76th Int. Workshop, Puchberg/Schneeberg (Austria). Physica Verlag, Heidelberg, pp. 13–20.

Blum, A., Hopcroft, J., Kannan, R., 2016. Foundations of Data Science. `http://www.cs.cornell.edu/jeh/bookJan25_2016.pdf`.

Cardot, H., Cénac, P., Monnez, J.-M., 2012. A fast and recursive algorithm for clustering large datasets. Comput. Statist. Data Anal. 56 (6), 1434–1449.

Cohn, H., Kumar, A., Miller, S., Radchenko, D., Viazovska, M., 2017. The sphere packing problem in dimension 24. Annals of Mathematics 185 (3), 1017–1033.

Cortés, J., Bullo, F., 2005. Coordination and geometric optimization via distributed dynamical systems. SIAM Journal on Control and Optimization 44 (5), 1543–1574.

Cortés, J., Bullo, F., 2009. Nonsmooth coordination and geometric optimization via distributed dynamical systems. SIAM Review 51 (1), 163–189.

Dobkin, D., Eppstein, D., 1993. Computing the discrepancy. In: Proceedings of the 9th Annual Symposium on Computational Geometry. ACM, pp. 47–52.

References II

Du, Q., Faber, V., Gunzburger, M., 1999. Centroidal Voronoi tessellations: applications and algorithms. SIAM Review 41 (4), 637–676.

Fang, K.-T., Lu, X., Winker, P., 2003. Lower bounds for centered and wrap-around L_2-discrepancies and construction of uniform designs by threshold accepting. Journal of Complexity 19 (5), 692–711.

Fang, K.-T., Ma, C.-X., 2001. Wrap-around L_2-discrepancy of random sampling, Latin hypercube and uniform designs. Journal of Complexity 17 (4), 608–624.

Fang, K.-T., Tang, Y., Yin, J., 2005. Lower bounds for wrap-around L_2-discrepancy and constructions of symmetrical uniform designs. Journal of Complexity 21 (5), 757–771.

Faure, H., 1982. Discrépances de suites associées à un système de numération (en dimension s). Acta Arithmetica 41, 337–351.

Franco, J., 2008. Planification d'expériences numériques en phase exploratoire pour la simulation de phénomènes complexes. Ph.D. Thesis, École Nationale Supérieure des Mines de Saint Étienne.

Franco, J., Vasseur, O., Corre, B., Sergent, M., 2009. Minimum Spanning Tree: a new approach to assess the quality of the design of computer experiments. Chemometrics and Intelligent Laboratory Systems 97, 164–169.

Gnewuch, M., Wahlström, M., Winzen, C., 2012. A new randomized algorithm to approximate the star discrepancy based on threshold accepting. SIAM Journal on Numerical Analysis 50 (2), 781–807.

Gonzalez, T., 1985. Clustering to minimize the maximum intercluster distance. Theoretical Computer Science 38, 293–306.

Guyader, A., Hengartner, N., Matzner-Løber, E., 2011. Simulation and estimation of extreme quantiles and extreme probabilities. Applied Mathematics & Optimization 64 (2), 171–196.

References III

Hickernell, F., 1998a. A generalized discrepancy and quadrature error bound. Mathematics of Computation 67 (221), 299–322.

Hickernell, F., 1998b. Lattice rules: how well do they measure up? In: Helekalek, P., Larcher, G. (Eds.), Random and Quasi-Random Point Sets. Vol. 138 of Lecture Notes in Statist. Springer, New York, pp. 109–166.

Hlawka, E., 1961. Funktionen von beschränkter variation in der theorie der gleichverteilung. Annali di Matematica Pura ed Applicata 54 (1), 325–333.

Johnson, M., Moore, L., Ylvisaker, D., 1990. Minimax and maximin distance designs. Journal of Statistical Planning and Inference 26, 131–148.

Joseph, V., Gul, E., Ba, S., 2015. Maximum projection designs for computer experiments. Biometrika 102 (2), 371–380.

Jourdan, A., Franco, J., 2010. Optimal Latin hypercube designs for the Kullback–Leibler criterion. AStA Advances in Statistical Analysis 94 (4), 341–351.

Koksma, J., 1942/1943. Een algemeene stelling uit de theorie der gelijkmatige verdeeling modulo 1. Mathematica B (Zutphen) 11, 7–11.

Korobov, N., 1960. Properties and calculation of optimal coefficients. Doklady Akademii Nauk SSSR 132 (5), 1009–1012.

Kuipers, L., Niederreiter, H., 1974. Uniform Distribution of Sequences. Wiley, New York.

Landkof, N., 1972. Foundations of Modern Potential Theory. Springer, Berlin.

Lekivetz, R., Jones, B., 2015. Fast flexible space-filling designs for nonrectangular regions. Quality and Reliability Engineering International 31 (5), 829–837.

References IV

Lloyd, S., 1982. Least squares quantization in PCM. IEEE Transactions on Information Theory 28 (2), 129–137.

Lubachevsky, B., 1991. How to simulate billiards and similar systems. Journal of Computational Physics 94 (2), 255–283.

Lubachevsky, B., Stillinger, F., 1990. Geometric properties of random disk packings. Journal of Statistical Physics 60 (5-6), 561–583.

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: LeCam, L., Neyman, J. (Eds.), Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Vol. 1. pp. 281–297.

Maisonneuve, D., 1972. Recherche et utilisation des "bons treillis". programmation et résultats numériques. In: Zaremba, S. (Ed.), Applications of Number Theory to Numerical Analysis. Academic Press, New York, pp. 121–201.

Mak, S., Joseph, V., 2016. Minimax and minimax projection designs using clustering. arXiv:1602.03938v3.

McKay, M., Beckman, R., Conover, W., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21 (2), 239Ű–245.

Morris, M., Mitchell, T., 1995. Exploratory designs for computational experiments. Journal of Statistical Planning and Inference 43, 381–402.

Müller, W., 2007. Collecting Spatial Data. Springer, Berlin, [3rd ed.].

Niederreiter, H., 1992. Random Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia.

Nuyens, D., 2007. Fast construction of good lattice rules. Ph.D. Thesis, Katholieke Univ. Leuven.

Oler, N., 1961. A finite packing problem. Canadian Mathematical Bulletin 4, 153–155.

References V

Owen, A., 1995. Randomly permuted (t, m, s)-nets and (t, s)-sequences. In: Niederreiter, H., Shiue, P. J.-S. (Eds.), Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing. Springer, New York, pp. 299–317.

Owen, A., 1998. Latin supercube sampling for very high-dimensional simulations. ACM Transactions on Modeling and Computer Simulation 8 (1), 71–102.

Penrose, M., Yukich, J., 2003. Weak laws of large numbers in geometric probability. The Annals of Applied Probability 13 (1), 277–303.

Penrose, M., Yukich, J., 2011. Laws of large numbers and nearest neighbor distances. In: Wells, M., Sengupta, A. (Eds.), Advances in Directional and Linear Statistics. A Festschrift for Sreenivasa Rao Jammalamadaka. Springer, pp. 189–199.

Pronzato, L., 2017a. Minimax and maximin space-filling designs: some properties and methods for construction. Journal de la Société Française de Statistique 158 (1), 7–36.

Pronzato, L., 2017b. On the construction of minimax-distance (sub-)optimal designs. A workshop on Latest Advances in the Theory and Applications of Design and Analysis of Experiments, Banff International Research Station for Mathematical Innovation and Discovery (BIRS), Canada, August 6-11, http://www.birs.ca/events/2017/5-day-workshops/17w5007/videos/watch/201708110847-Pronzato.html.

Pronzato, L., 2017c. On the elimination of inessential points in the smallest enclosing ball problem. Optimization Methods and SoftwareTo appear.

Pronzato, L., Müller, W., 2012. Design of computer experiments: space filling and beyond. Statistics and Computing 22, 681–701.

Pronzato, L., Wynn, H., Zhigljavsky, A., 2016. Extremal measures maximizing functionals based on simplicial volumes. Statistical Papers 57 (4), 1059–1075, hal-01308116.

References VI

Redmond, C., Yukich, J., 1994. Limit theorems and rates of convergence for euclidean functionals. The Annals of Applied Probability 4 (4), 1057–1073.

Redmond, C., Yukich, J., 1996. Asymptotics for Euclidian functionals with power-weighted edges. Stochastic Processes and their Applications 61, 289–304.

Riccomagno, E., Schwabe, R., Wynn, H., 1997. Lattice-based D-optimum design for Fourier regression. Annals of Statistics 25 (6), 2313–2327.

Saff, E., 2010. Logarithmic potential theory with applications to approximation theory. Surveys in Approximation Theory 5 (14), 165–200.

Sloan, I., Reztsov, A., 2002. Component-by-component construction of good lattice rules. Mathematics of Computation 71 (237), 263–273.

Sloan, I., Walsh, L., 1990. A computer search of rank-2 lattice rules for multidimensional quadrature. Mathematics of Computation 54 (189), 281–302.

Sobol', I., 1967. On the distribution of points in a cube and the approximate evaluation of integrals. USSR Computational Mathematics and Mathematical Physics 7 (4), 86–112.

Steele, J., 1981. Subadditive Euclidean functionals and nonlinear growth in geometric probability. The Annals of Probability, 365–376.

Thiémard, E., 2001. An algorithm to compute bounds for the star discrepancy. Journal of Complexity 17 (4), 850–880.

van Dam, E., 2008. Two-dimensional minimax Latine hypercube designs. Discrete Applied Math. 156 (18), 3483–3493.

van Dam, E., Hussage, B., den Hertog, D., Melissen, H., 2007. Maximin Latine hypercube designs in two dimensions. Operations Research 55 (1), 158–169.

References VII

van der Corput, J., 1935. Verteilungsfunktionen. I. Mitt. Proc. Akad. Wet. Amsterdam 38, 813–821.

Viana, F., 2013. Things you wanted to know about the Latin hypercube design and were afraid to ask. In: 10th World Congress on Structural and Multidisciplinary Optimization, Orlando, Florida, USA.

Viazovska, M., 2016. The sphere packing problem in dimension 8. arXiv preprint arXiv:1603.04246.

Wade, A., 2007. Explicit laws of large numbers for random nearest-neighbour-type graphs. Advances in Applied Probability 39 (2), 326–342.

Wahl, F., Mercadier, C., Helbert, C., 2014. Measuring the quality of maximin space-filling designs. Hal-00955294.

Yildirim, E., 2008. Two algorithms for the minimum enclosing ball problem. SIAM Journal on Optimization 19 (3), 1368–1391.

Yukich, J., 1998. Probability Theory of Classical Euclidean Optimization Problems. Springer, Berlin.

Zhigljavsky, A., Hamilton, E., 2010. Stopping rules in k-adaptive global random search algorithms. Journal of Global Optimization 48 (1), 87–97.

Zhigljavsky, A., Žilinskas, A., 2007. Stochastic Global Optimization. Springer, New York.