

Plans d'expériences numériques — (1) sans modèle —

LUC PRONZATO

Laboratoire I3S, CNRS-Univ. Nice Sophia Antipolis, France

Pec'num²⁰¹⁵

Objectifs

Expériences numériques : à base de simulations

- Habituellement, $\mathbf{x} \in \mathbb{R}^d \mapsto$ observation $Y(\mathbf{x})$ (phénomène physique)
- ici, simulation numérique : $Y(\mathbf{x}) = f(\mathbf{x})$, observation = évaluation d'une fonction $f(\cdot)$ inconnue
(pas de bruit de mesure)

Objectifs

Expériences numériques : à base de simulations

- Habituellement, $\mathbf{x} \in \mathbb{R}^d \mapsto$ observation $Y(\mathbf{x})$ (phénomène physique)
- ici, simulation numérique : $Y(\mathbf{x}) = f(\mathbf{x})$, observation = évaluation d'une fonction $f(\cdot)$ inconnue
(pas de bruit de mesure)

A partir de paires $(\mathbf{x}_i, f(\mathbf{x}_i))$, $i = 1, 2, \dots, n$

- optimisation : trouver $\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x})$
- inversion : reconstruire $\{\mathbf{x} \in \mathcal{X} : f(\mathbf{x}) = T\}$
- estimation d'une probabilité de défaillance : $\text{Prob}\{f(\mathbf{x}) > C\}$ quand $\mathbf{x}$ est distribué suivant une densité de probabilité $\phi(\cdot)$
- analyse de sensibilité
- **approximation/interpolation de $f(\cdot)$** par une fonction $\eta_n(\cdot)$, à construire

Objectif = approximation/interpolation

$f(\mathbf{x})$ une fonction inconnue, définie sur $\mathcal{X} \subset \mathbb{R}^d$

construire une “bonne” approximation $\eta_n(\cdot)$ de $f(\cdot)$ sur $\mathcal{X}$ à partir de paires $(\mathbf{x}_i, f(\mathbf{x}_i))$, $i = 1, 2, \dots, n$ (n n'est pas forcément fixé au départ)

Objectif = approximation/interpolation

$f(\mathbf{x})$ une fonction inconnue, définie sur $\mathcal{X} \subset \mathbb{R}^d$

construire une “bonne” approximation $\eta_n(\cdot)$ de $f(\cdot)$ sur $\mathcal{X}$ à partir de paires $(\mathbf{x}_i, f(\mathbf{x}_i))$, $i = 1, 2, \dots, n$ (n n'est pas forcément fixé au départ)

⇒ Puisque $f(\cdot)$ est inconnue, il faut observer partout !

⇒ disperser au maximum les n points $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ dans $\mathcal{X}$

(l'uniformité peut être justifiée précisément (Biedermann and Dette, 2001))

Objectif = approximation/interpolation

$f(\mathbf{x})$ une fonction inconnue, définie sur $\mathcal{X} \subset \mathbb{R}^d$

construire une “bonne” approximation $\eta_n(\cdot)$ de $f(\cdot)$ sur $\mathcal{X}$ à partir de paires $(\mathbf{x}_i, f(\mathbf{x}_i))$, $i = 1, 2, \dots, n$ (n n'est pas forcément fixé au départ)

⇒ Puisque $f(\cdot)$ est inconnue, il faut observer partout !

⇒ disperser au maximum les n points $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ dans $\mathcal{X}$

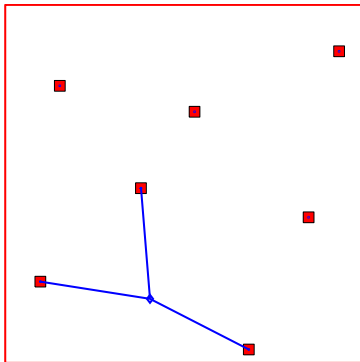
(l'uniformité peut être justifiée précisément (Biedermann and Dette, 2001))

▷ $\mathbf{X}_n$ est le plan d'expérience (à n points) ◁

Que veut dire observer partout, disperser $\mathbf{X}_n$ au maximum ?

Aperçu général : 3 familles de critères

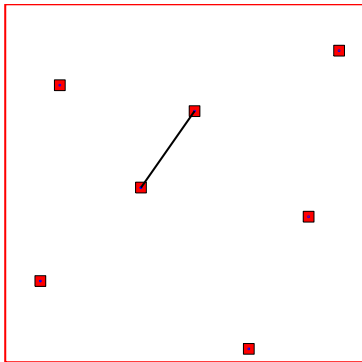
1. Inter-distance : entre $\mathbf{X}_n$ et $\mathcal{X}$ (miniMax, dispersion)



... assez compliqué

Aperçu général : 3 familles de critères

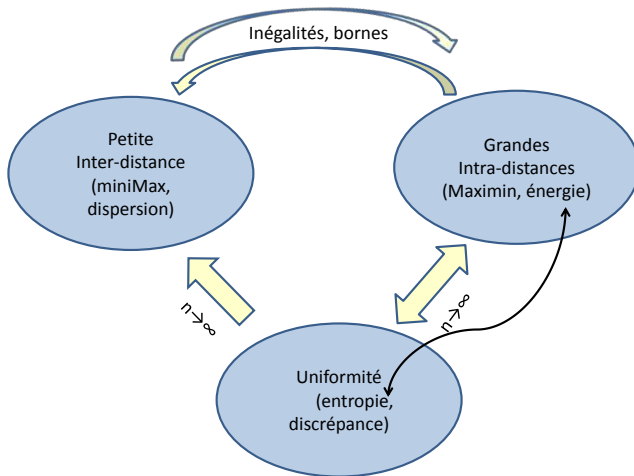
1. Inter-distance : entre $\mathbf{X}_n$ et $\mathcal{X}$ (miniMax, dispersion)
2. Intra-distances : au sein de $\mathbf{X}_n$, entre les points $\mathbf{x}_i$, $i = 1, \dots, n$
(Maximin, énergie...)



... plus facile

Aperçu général : 3 familles de critères

1. Inter-distance : entre $\mathbf{X}_n$ et $\mathcal{X}$ (miniMax, dispersion)
2. Intra-distances : au sein de $\mathbf{X}_n$, entre les points $\mathbf{x}_i$, $i = 1, \dots, n$
(Maximin, énergie...)
3. Uniformité de la distribution des $\mathbf{x}_i$, $i = 1, \dots, n$
(entropie, discrédance $\Rightarrow$ générer des suites infinies de points)



Plan

- 1 Critères géométriques *space-filling*
 - 1.1 miniMax & Maximin : généralités
 - 1.2 Hypercubes Latins
 - 1.3 Critère miniMax (inter-distance)
 - 1.4 Critère Maximin (intra-distances)
 - 1.5 Maximin régularisé, énergie
- 2 Uniformité: quasi Monte-Carlo, discrédance
 - 2.1 Entropie, graphes optimaux
 - 2.2 Discrédance : motivation
 - 2.3 Critères de discrédance
 - 2.4 Suites à faible discrédance
 - 2.5 (t, m, d) -nets & suites- (t, d)
- 3 Critère miniMax & dispersion
 - 3.1 Dispersion
 - 3.2 Suites à faible dispersion
- 4 Conclusions partie (1)

1 Critères géométriques *space-filling*

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

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

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

⇒ **Inter-distance** entre $\mathbf{X}_n$ et $\mathcal{X}$

$$\begin{aligned} d = 1 &\Leftrightarrow x_i = (2i - 1)/(2n), \quad i = 1, \dots, n \\ &\Rightarrow \Phi_{mM}^* = 1/(2n) \end{aligned}$$

$d > 1 \Leftrightarrow$ recouvrement par des sphères (*sphere-covering*)

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

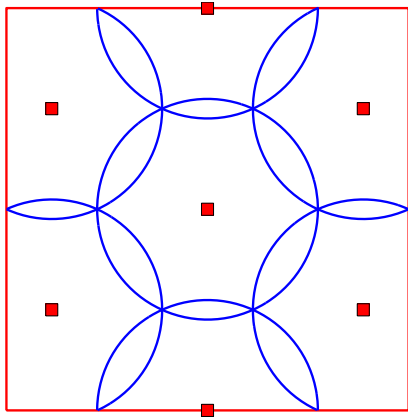
▣ **Intra-distances** pour $\mathbf{X}_n$ (entre points $\mathbf{x}_i$ de $\mathbf{X}_n$)

$$d = 1 \Leftrightarrow x_i = (i - 1)/(n - 1), \quad i = 1, \dots, n$$
$$\Rightarrow \Phi_{Mm}^* = 1/(n - 1)$$

$d > 1 \Leftrightarrow$ empilement de sphères (*sphere-packing*)

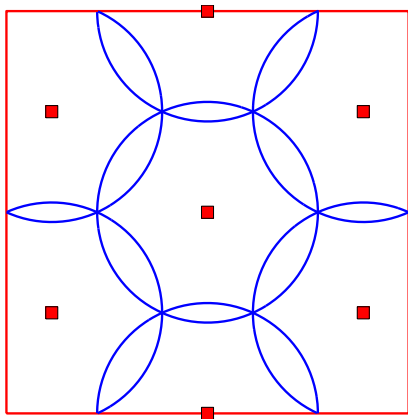
Exemples :

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

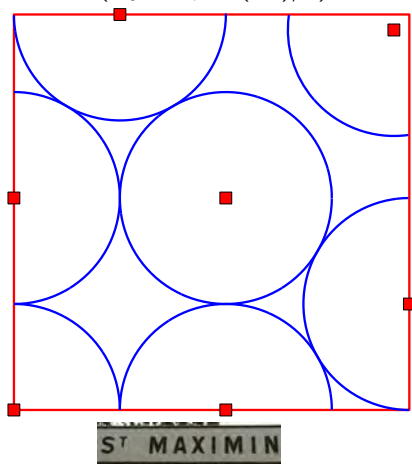


Exemples :

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

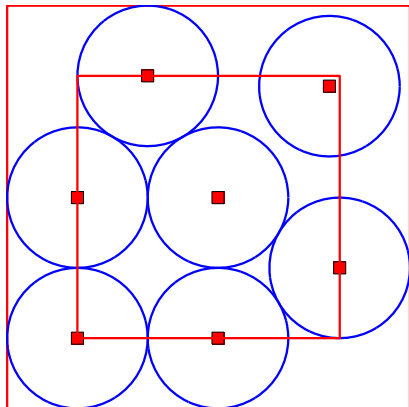


② Maximin $d = 2, n = 7$
(rayon = $\phi_{Mm}(\mathbf{X}_n)/2$)



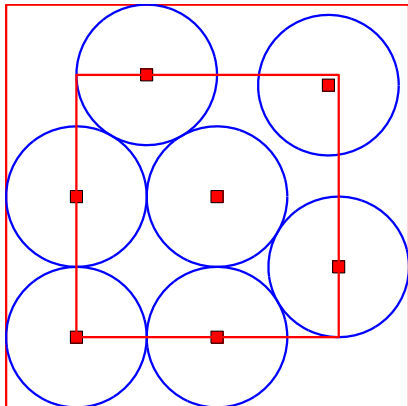
Pourquoi Maximin $\Leftrightarrow$ empilement de sphères ?

$$d = 2, n = 7$$

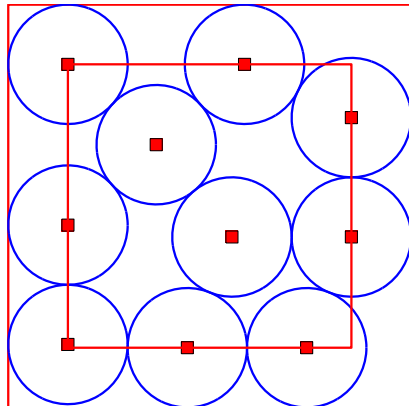


Pourquoi Maximin $\Leftrightarrow$ empilement de sphères ?

$d = 2, n = 7$



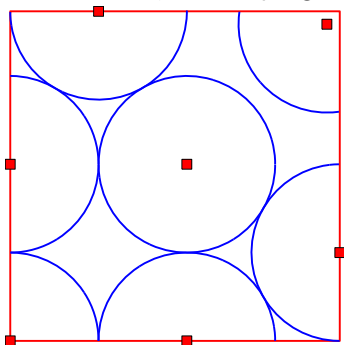
$d = 2, n = 10$



Quelques difficultés :

a) Optimum local ?

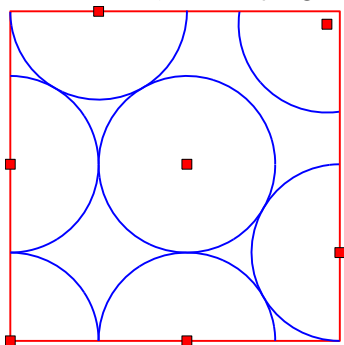
Maximin $d = 2, n = 7$, opt. global



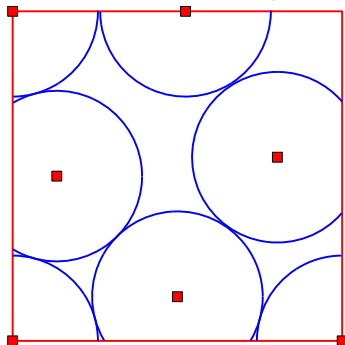
Quelques difficultés :

a) Optimum local ?

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



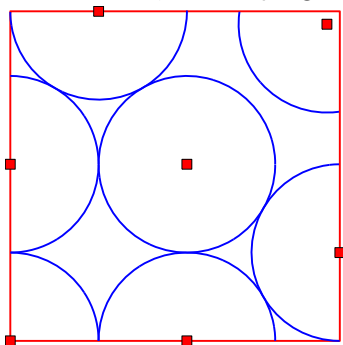
Maximin $d = 2, n = 7$, opt. local



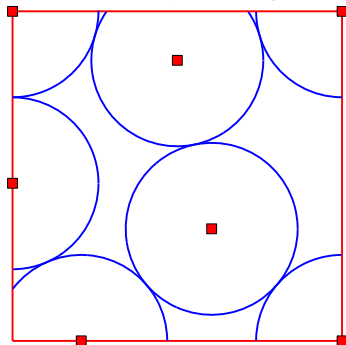
Quelques difficultés :

a) Optimum local ?

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



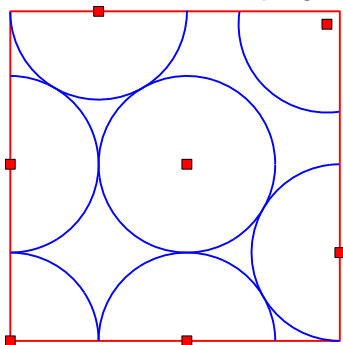
Maximin $d = 2, n = 7$, opt. local



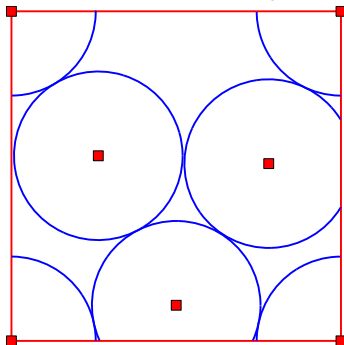
Quelques difficultés :

a) Optimum local ?

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



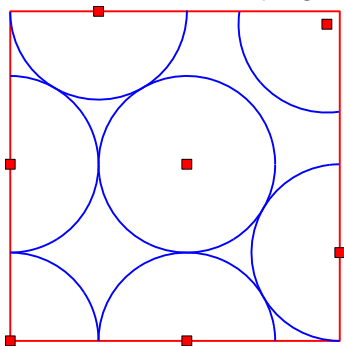
Maximin $d = 2, n = 7$, opt. local



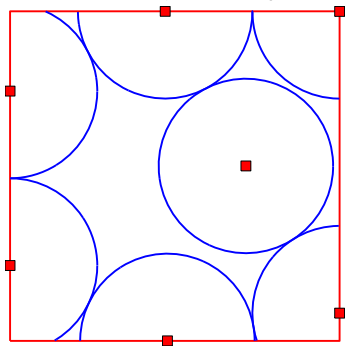
Quelques difficultés :

a) Optimum local ?

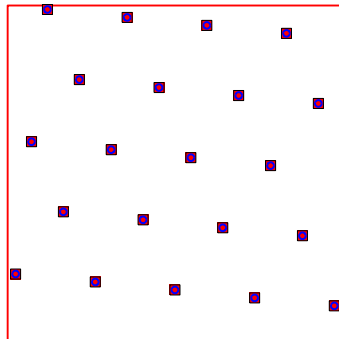
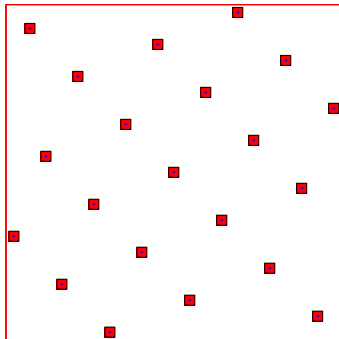
Maximin $d = 2, n = 7$, opt. global

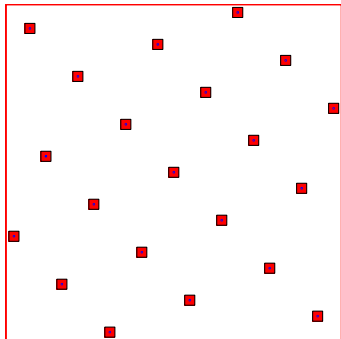


Maximin $d = 2, n = 7$, opt. local

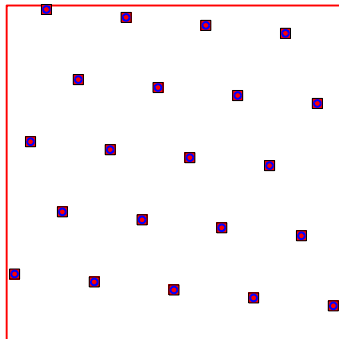


b) Intuition trompeuse ?



b) Intuition trompeuse ?

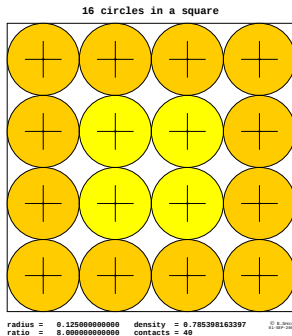
Géométrie : $\Phi_{Mm}(\mathbf{X}_n) = 0.2020$
 $\Phi_{mM}(\mathbf{X}_n) = 0.2357$
 Uniformité : $D_{Cent, L_2}(\mathbf{X}_n) = 0.0280$
 $D_{WA, L_2}(\mathbf{X}_n) = 0.0388$



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

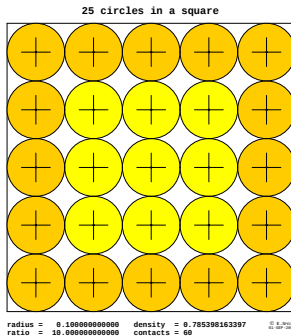
c) Empilement de sphères : pas de solution triviale,
voir <http://www.packomania.com/>

$$d = 2, n = 16$$



c) Empilement de sphères : pas de solution triviale,
voir <http://www.packomania.com/>

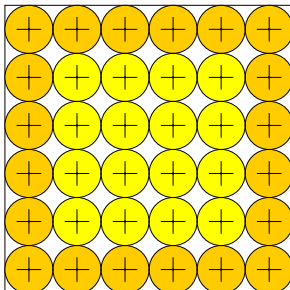
$$d = 2, n = 25$$



c) Empilement de sphères : pas de solution triviale,
voir <http://www.packomania.com/>

$$d = 2, n = 36$$

36 circles in a square



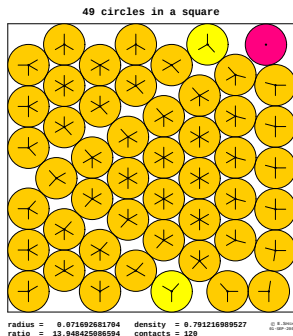
```
radius = 0.08333333333333333
ratio = 12.000000000000000
```

```
density = 0.785398163397
contacts = 84
```

© S. SPOHN
02-00P-2000

c) Empilement de sphères : pas de solution triviale,
voir <http://www.packomania.com/>

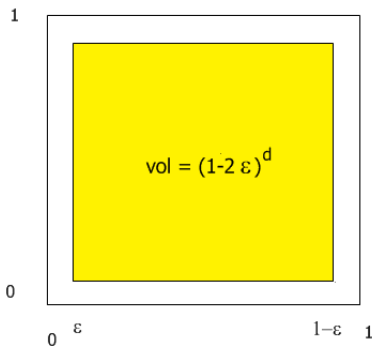
$$d = 2, n = 49$$



Il est plus facile d'empiler des cubes ! (voir § 3.1)

d) Fléau de la dimension : quand $d \rightarrow \infty$, tout le volume est sur les bords

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

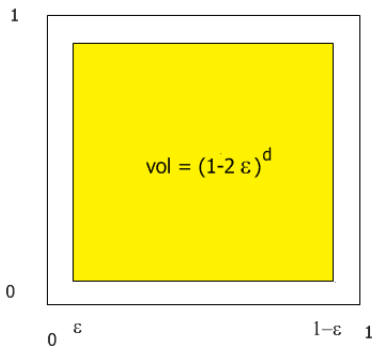


Pour ϵ fixé, volume partie centrale $= (1 - 2\epsilon)^d \rightarrow 0$ quand $d \rightarrow \infty$

Ex : 64 boules dans un cube, en 4 couches régulières de 16 boules
→ 60 sont sur les bords !

d) Fléau de la dimension : quand $d \rightarrow \infty$, tout le volume est sur les bords

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



Pour ϵ fixé, volume partie centrale $= (1 - 2\epsilon)^d \rightarrow 0$ quand $d \rightarrow \infty$

Ex : 64 boules dans un cube, en 4 couches régulières de 16 boules
 $\rightarrow$ 60 sont sur les bords !

On va réduire nos ambitions : l'optimum étant difficile à atteindre, essayer seulement de trouver des plans "raisonnables"

1.2 Hypercubes Latins (*Latin hypercubes*)

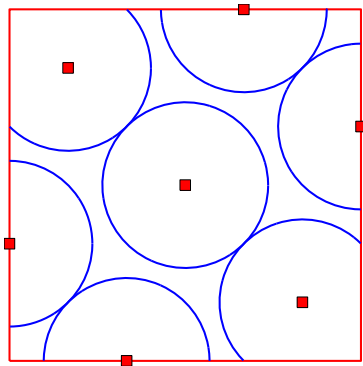
Objectif : assurer de bonnes propriétés de projection sur chaque axe

chaque projection 1D est Maximin-optimale

$\{\mathbf{x}_i\}_\ell \in \{0, \frac{1}{n-1}, \dots, \frac{k-1}{n-1}, \dots, 1\}$ pour tout $\ell = 1, \dots, d$

⇒ ensemble fini de $(n!)^{d-1}$ plans possibles

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



1.2 Hypercubes Latins (*Latin hypercubes*)

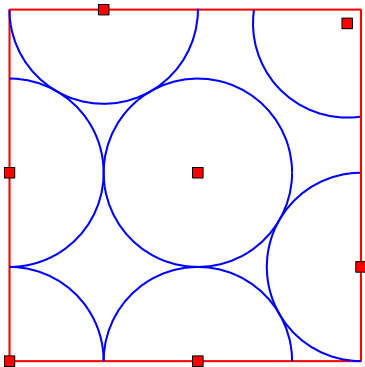
Objectif : assurer de bonnes propriétés de projection sur chaque axe

chaque projection 1D est Maximin-optimale

$\{\mathbf{x}_i\}_\ell \in \{0, \frac{1}{n-1}, \dots, \frac{k-1}{n-1}, \dots, 1\}$ pour tout $\ell = 1, \dots, d$

⇒ ensemble fini de $(n!)^{d-1}$ plans possibles

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



1.2 Hypercubes Latins (*Latin hypercubes*)

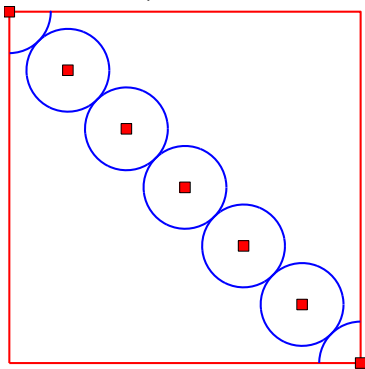
Objectif : assurer de bonnes propriétés de projection sur chaque axe

chaque projection 1D est Maximin-optimale

$\{\mathbf{x}_i\}_\ell \in \{0, \frac{1}{n-1}, \dots, \frac{k-1}{n-1}, \dots, 1\}$ pour tout $\ell = 1, \dots, d$

⇒ ensemble fini de $(n!)^{d-1}$ plans possibles

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



Un plan Lh a seulement de bonnes propriétés de projection 1D !

- ▣ Il faut encore optimiser une autre caractéristique de dispersion en dimension supérieure (typiquement recuit simulé, mais autres heuristiques envisageables)

Un plan Lh a seulement de bonnes propriétés de projection 1D !

⇒ Il faut encore optimiser une autre caractéristique de dispersion en dimension supérieure (typiquement recuit simulé, mais autres heuristiques envisageables)

⇒ Travailler dans la classe des plans Lh permet alors d'assurer de bonnes propriétés de projection 1D

Important quand $f(\cdot)$ peut ne pas dépendre de certains facteurs $\{\mathbf{x}\}_\ell$:

- pas de répétitions de points quand on supprime une composante
- la projection sur $d' < d$ composantes est encore un Lh (mais pas forcément avec bonne distribution des points si $d' > 1$)

Littérature abondante depuis (McKay et al., 1979), voir (Viana, 2013)

La contrainte Lh fait perdre un peu en optimalité

Optimisation dans la classe des plans Lh :

Soit $\mathbf{X}_n$ un plan Lh à n points:

- choisir une coordonnée ℓ parmi les d
- choisir une paire de points $\mathbf{x}_i$ et $\mathbf{x}_j$ de $\mathbf{X}_n$
- échanger leurs coordonnées numéro ℓ
- $\Rightarrow \mathbf{X}_n^+$, qui est encore un plan Lh (avec $dn(n-1)/2$ constructions possibles)

(on peut aussi échanger plusieurs paires de points à la fois)

Optimisation dans la classe des plans Lh :

Soit $\mathbf{X}_n$ un plan Lh à n points:

- choisir une coordonnée ℓ parmi les d
- choisir une paire de points $\mathbf{x}_i$ et $\mathbf{x}_j$ de $\mathbf{X}_n$
- échanger leurs coordonnées numéro ℓ
- $\Rightarrow \mathbf{X}_n^+$, qui est encore un plan Lh (avec $dn(n-1)/2$ constructions possibles)

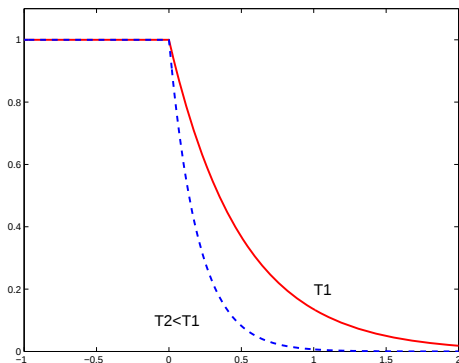
(on peut aussi échanger plusieurs paires de points à la fois)

Recuit simulé (minimisation de $\Phi(\cdot)$) — principe :

- 0) partant du plan Lh $\mathbf{X}_n^0$, faire $k = 0$
- 1) générer le plan Lh $\mathbf{X}_n^{k+}$ à partir de $\mathbf{X}_n^k$
- 2) calculer $\Delta\Phi_k = \Phi(\mathbf{X}_n^{k+}) - \Phi(\mathbf{X}_n^k)$
- 3) Accepter $\mathbf{X}_n^{k+}$, c'est-à-dire faire $\mathbf{X}_n^{k+1} = \mathbf{X}_n^{k+}$ avec la probabilité $P_k = \min \left\{ 1, \exp \left(-\frac{\Delta\Phi_k}{T_k} \right) \right\}$, garder $\mathbf{X}_n^{k+1} = \mathbf{X}_n^k$ avec la prob. $1 - P_k$
 $k \leftarrow k + 1$, retourner en 1

X_n^{k+} tel que $\Delta\Phi_k < 0$ est toujours accepté

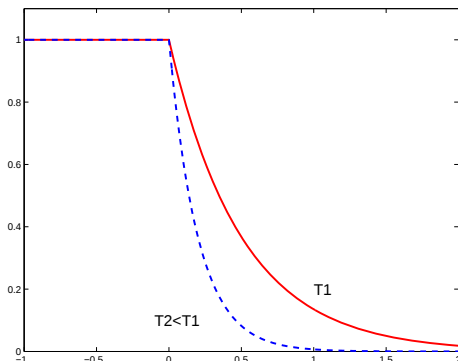
X_n^{k+} tel que $\Delta\Phi_k > 0$ est plus facilement accepté pour T_1 que pour $T_2 < T_1$



Prendre T_0 suffisamment grand (pour sortir des minima locaux), puis faire décroître T_k (par ex., $T_k = \frac{T_0}{\log(k+1)}$, ou $T_k = \alpha^k T_0$ avec $\alpha < 1$)

$\mathbf{X}_n^{k+}$ tel que $\Delta\Phi_k < 0$ est toujours accepté

$\mathbf{X}_n^{k+}$ tel que $\Delta\Phi_k > 0$ est plus facilement accepté pour T_1 que pour $T_2 < T_1$



⇒ Prendre T_0 suffisamment grand (pour sortir des minima locaux), puis faire décroître T_k (par ex., $T_k = \frac{T_0}{\log(k+1)}$, ou $T_k = \alpha^k T_0$ avec $\alpha < 1$)

Beaucoup de variantes, littérature abondante. . .

Toujours garder en mémoire le meilleur $\mathbf{X}_n^k$ trouvé le long de la trajectoire de l'algorithme !!!

⇒ Garantit la convergence vers l'optimum quand $k \rightarrow \infty$ sous des conditions très générales

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

Φ_{mM} est intéressant en terme d'approximation:

tout $\mathbf{x}$ de $\mathcal{X}$ à distance $< \Phi_{mM}$ d'un point $\mathbf{x}_i$!

Au moins 3 possibilités pour évaluer $\Phi_{mM}(\mathbf{X}_n)$:

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

Φ_{mM} est intéressant en terme d'approximation:

tout $\mathbf{x}$ de $\mathcal{X}$ à distance $< \Phi_{mM}$ d'un point $\mathbf{x}_i$!

Au moins 3 possibilités pour évaluer $\Phi_{mM}(\mathbf{X}_n)$:

A) Remplacer $\mathcal{X}$ par un ensemble fini $\mathcal{X}_Q$ à Q points

(une grille régulière, Q éléments d'une suite à faible discrédance):

$$\Rightarrow \Phi_{mM}(\mathbf{X}_n) \simeq \max_{\mathbf{x}^{(k)} \in \mathcal{X}_Q} \min_i \|\mathbf{x}^{(k)} - \mathbf{x}_i\|$$

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

Φ_{mM} est intéressant en terme d'approximation:

tout $\mathbf{x}$ de $\mathcal{X}$ à distance $< \Phi_{mM}$ d'un point $\mathbf{x}_i$!

Au moins 3 possibilités pour évaluer $\Phi_{mM}(\mathbf{X}_n)$:

A) Remplacer $\mathcal{X}$ par un ensemble fini $\mathcal{X}_Q$ à Q points

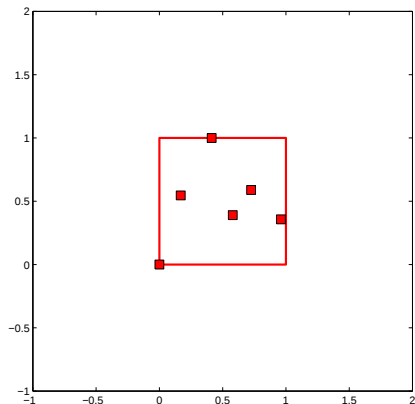
(une grille régulière, Q éléments d'une suite à faible discrédance):

$$\Rightarrow \Phi_{mM}(\mathbf{X}_n) \simeq \max_{\mathbf{x}^{(k)} \in \mathcal{X}_Q} \min_i \|\mathbf{x}^{(k)} - \mathbf{x}_i\|$$

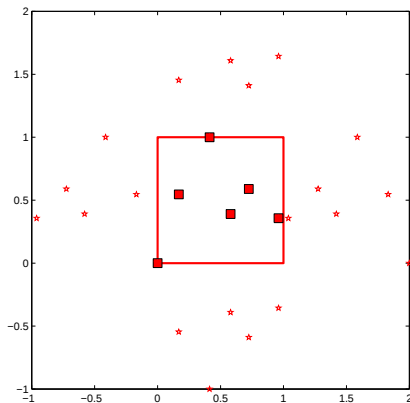
B) Triangulation (tessellation) de Delaunay (**quand $\mathcal{X} = \text{hypercube}$**) (Pronzato and Müller, 2012)

- $\mathbf{X}_n$ ($= n$ points dans $\mathcal{X} = \mathbb{I}^d = [0, 1]^d$), prendre tous les points symétriques par rapport aux $2d$ faces de dimension $d - 1$
 $\Rightarrow M = (2d + 1)n$ points
- Calculer la triangulation (tessellation) de Delaunay des M points
- Les points candidats pour $\arg \max_{\mathbf{x} \in \mathcal{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$ sont aux centres des sphères circonscrites

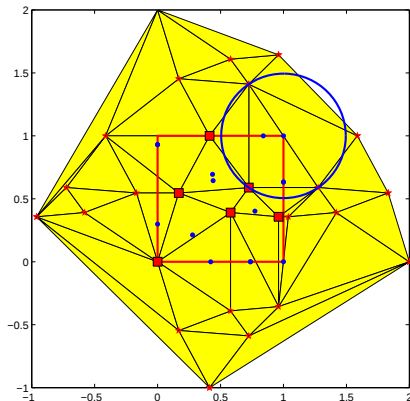
B) miniMax par triangulation de Delaunay



B) miniMax par triangulation de Delaunay

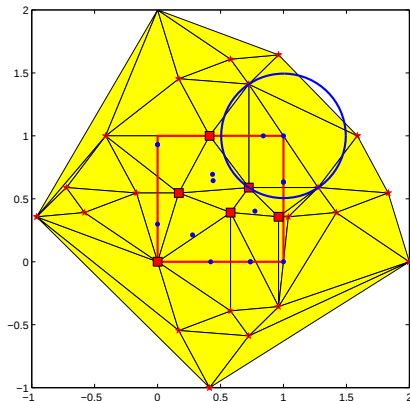


B) miniMax par triangulation de Delaunay



$n = 6$ points, 45 triangles, 15 cercles, le + grand est tracé

B) miniMax par triangulation de Delaunay



$n = 6$ points, 45 triangles, 15 cercles, le + grand est tracé

Le calcul de $\Phi_{mM}(\mathbf{X}_n)$ demeure coûteux: jusqu'à $M^{\lceil d/2 \rceil}$ simplexes (et autant de sphères circonscrites), avec $M = (2d + 1)n$

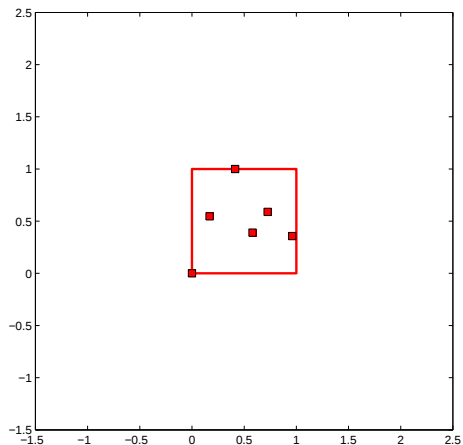
⇒ temps de calcul = $\mathcal{O}(M \log M + M^{\lceil d/2 \rceil})$

C) Diagramme de Voronoï (**quand $\mathcal{X}$ = un polytope quelconque (polyèdre compact) $\subset \mathbb{I}^d = [0, 1]^d$**)

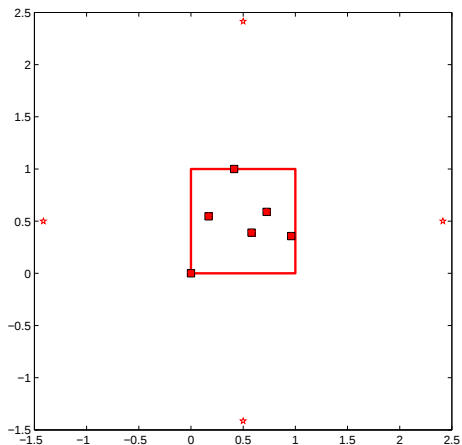
C) Diagramme de Voronoï (**quand $\mathcal{X}$ = un polytope quelconque (polyèdre compact) $\subset \mathbb{I}^d = [0, 1]^d$**)

- $\mathbf{X}_n$ (= n points in $\mathcal{X} = [0, 1]^d$)
 $\cup \{\text{centre } \pm \alpha \mathbf{e}_i, i = 1, \dots, d\}$ (avec $\alpha \geq \sqrt{d} + 1/2$)
 $\Rightarrow M = n + 2d$ points
- Calculer le diagramme de Voronoï
- Intersecter toutes les cellules $\mathcal{V}_i$ (polyédriques) avec $\mathcal{X}$ (également polyédrique)
- Les points candidats pour $\arg \max_{\mathbf{x} \in \mathcal{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$ sont des sommets de la partition de $\mathcal{X}$

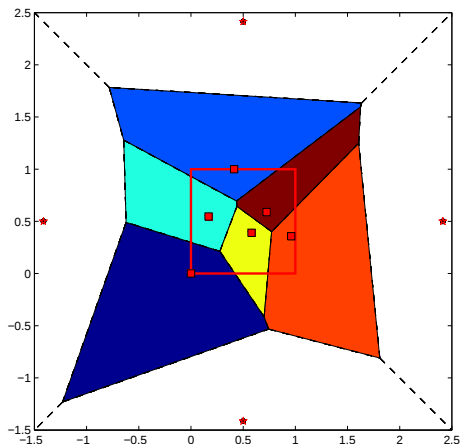
C) miniMax par diagramme de Voronoï



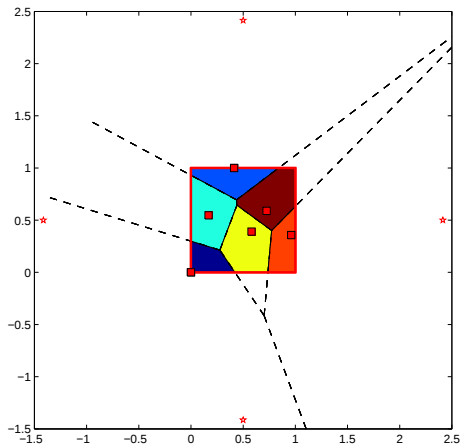
C) miniMax par diagramme de Voronoï



C) miniMax par diagramme de Voronoï

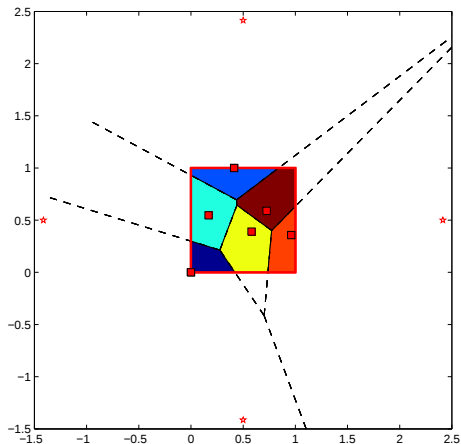


C) miniMax par diagramme de Voronoï



$n = 6$ points, 6 cellules, 28 sommets $\mathbf{x}^{(k)}$ testés pour $\min_i \|\mathbf{x}^{(k)} - \mathbf{x}_i\|$

C) miniMax par diagramme de Voronoï



$n = 6$ points, 6 cellules, 28 sommets $\mathbf{x}^{(k)}$ testés pour $\min_i \|\mathbf{x}^{(k)} - \mathbf{x}_i\|$

Le calcul de $\Phi_{mM}(\mathbf{X}_n)$ demeure coûteux:

jusqu'à $M(M-1)/2$ faces et $2M^{\lceil d/2 \rceil} / \lceil d/2 \rceil!$ sommets, avec $M = n + 2d$

Minimisation de $\Phi_{mM}(\mathbf{X}_n) = \max_{\mathbf{x} \in \mathcal{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$?

Difficulté: la fonction $\mathbf{X}_n \mapsto \min_i \|\mathbf{x} - \mathbf{x}_i\|$ est ni convexe, ni différentiable !

Si n et d pas trop grands et $\mathcal{X}$ assez simple, on sait calculer $\Phi_{mM}(\mathbf{X}_n)$ (outils de géométrie algorithmique: Delaunay & Voronoï)

Minimisation de $\Phi_{mM}(\mathbf{X}_n) = \max_{\mathbf{x} \in \mathcal{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$?

Difficulté: la fonction $\mathbf{X}_n \mapsto \min_i \|\mathbf{x} - \mathbf{x}_i\|$ est ni convexe, ni différentiable !

Si n et d pas trop grands et $\mathcal{X}$ assez simple, on sait calculer $\Phi_{mM}(\mathbf{X}_n)$ (outils de géométrie algorithmique: Delaunay & Voronoï)

- ⇒ Utiliser un outil d'optimisation globale (ex., recuit simulé)
- ⇒ Descente locale avec initialisation ad'hoc
(par ex. aléatoire ⇒ *multistart*)

Minimisation de $\Phi_{mM}(\mathbf{X}_n) = \max_{\mathbf{x} \in \mathcal{X}} \min_i \|\mathbf{x} - \mathbf{x}_i\|$?

Difficulté: la fonction $\mathbf{X}_n \mapsto \min_i \|\mathbf{x} - \mathbf{x}_i\|$ est ni convexe, ni différentiable !

Si n et d pas trop grands et $\mathcal{X}$ assez simple, on sait calculer $\Phi_{mM}(\mathbf{X}_n)$ (outils de géométrie algorithmique: Delaunay & Voronoï)

- ⇒ Utiliser un outil d'optimisation globale (ex., recuit simulé)
- ⇒ Descente locale avec initialisation ad'hoc
(par ex. aléatoire ⇒ *multistart*)

Problème assez difficile, mais :

$\Phi_{mM}(\cdot)$ est globalement Lipschitz (avec constante 1) (Cortés and Bullo, 2005, 2009)

→ $\Phi_{mM}(\cdot)$ est différentiable presque partout

Algorithme de type gradient généralisé (voir (Clarke, 1975))

$$\mathbf{X}_n^{k+1} = \mathbf{X}_n^k - \gamma_k \partial \Phi_{mM;i,\mathbf{x}_*}(\mathbf{X}_n^k)$$

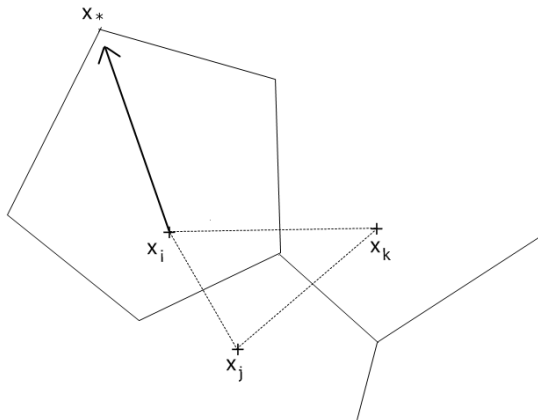
avec

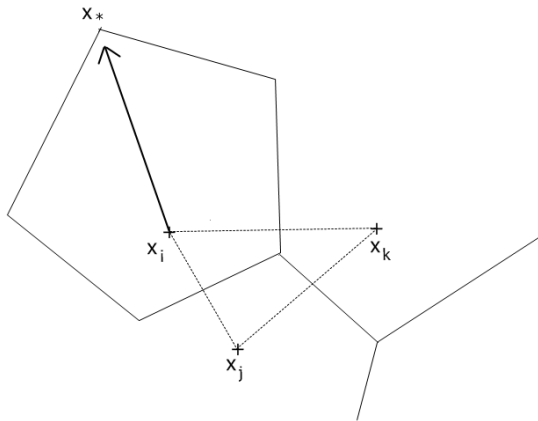
$$\gamma_k \searrow 0, \quad \sum_k \gamma_k = \infty$$

$$\partial \Phi_{mM;i,\mathbf{x}_*}(\mathbf{X}_n) = \frac{\partial \Phi_{mM;i,\mathbf{x}_*}(\mathbf{X}_n)}{\partial \mathbf{x}_i} = (\mathbf{0}, \dots, \mathbf{0}, \underbrace{\frac{\mathbf{x}_i - \mathbf{x}_*}{\|\mathbf{x}_i - \mathbf{x}_*\|}}_{i\text{ème position}}, \mathbf{0}, \dots, \mathbf{0}),$$

où

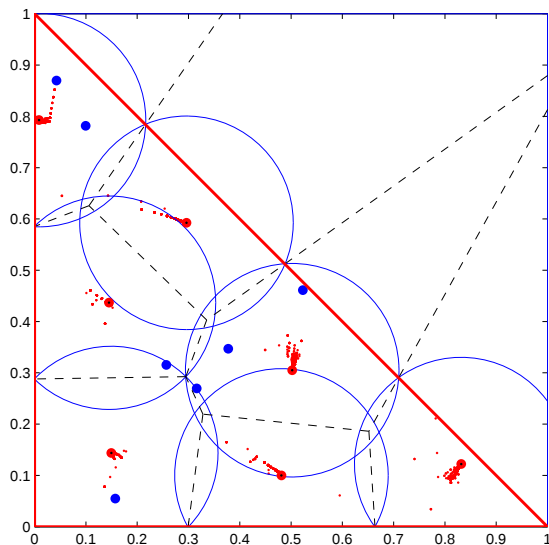
- $\Phi_{mM;i,\mathbf{x}_*}(\mathbf{X}_n) = \|\mathbf{x}_* - \mathbf{x}_i\|$
- $\mathbf{x}_*$ et i sont tels que $\|\mathbf{x}_* - \mathbf{x}_i\| = \Phi_{mM}(\mathbf{X}_n^k)$
 ($\mathbf{x}_*$ trouvé par triangulation de Delaunay ou diagramme de Voronoï)
 ($\mathbf{X}_n^k$ reste loin des bords de $\mathcal{X}$)





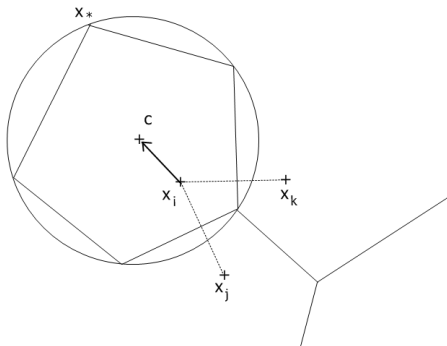
Peut être fait dans une seule cellule de Voronoï (celle d'indice i , telle que $\|x_* - x_i\| = \Phi_{mM}(X_n^k)$), ou dans toutes les cellules en parallèle (Cortés and Bullo, 2005, 2009) : calcul exact de gradient généralisé ($d = 2$) sans avoir $\gamma_k \searrow 0$ (évolution équa. diff. temps continu)

Ex: Minimisation locale de $\Phi_{mM}(\mathbf{X}_n)$, $n = 7$, $\mathcal{X} = \text{simplexe } 0 \leq x_1, 0 \leq x_2, x_1 + x_2 \leq 1$
 (rayon = $\Phi_{mM}(\mathbf{X}_n)$)



Quelques variations:

- 1) Remplacer $\mathbf{x}_i^k$ par $\mathbf{x}_i^{k+1}$ = centre sphère $\mathcal{S}_i^*$ circonscrite à la cellule $\mathcal{V}_i$, de volume minimal
- Voir (Cortés and Bullo, 2005, 2009) pour une version à temps continu
- $\mathcal{S}_i^*$ se calcule: voir Welzl (1991); Botkin and Turova-Botkina (1994); Xu et al. (2003)
- Converge vers un min. local de $\Phi_{mM}(\mathbf{X}_n)$



2) Remplacer $\mathbf{x}_i^k$ par $\mathbf{x}_i^{k+1} = \text{centre } \mathbf{c}_E(\mathcal{V}_i)$ de l'ellipsoïde $\mathcal{E}_i^*$ de vol. min. contenant $\mathcal{V}_i$ ($\mathcal{E}_i^* = \text{ellipsoïde de John (1985), se calcule}$)

⇒ Ne converge probablement pas vers un min. local de $\Phi_{mM}(\mathbf{X}_n)$

2) Remplacer $\mathbf{x}_i^k$ par $\mathbf{x}_i^{k+1} =$ centre $\mathbf{c}_E(\mathcal{V}_i)$ de l'ellipsoïde $\mathcal{E}_i^*$ de vol. min. contenant $\mathcal{V}_i$ ($\mathcal{E}_i^* =$ ellipsoïde de John (1985), se calcule)

⇒ Ne converge probablement pas vers un min. local de $\Phi_{mM}(\mathbf{X}_n)$

3) Remplacer $\mathbf{x}_i^k$ par $\mathbf{x}_i^{k+1} =$ centre de gravité $\mathbf{c}_G(\mathcal{V}_i)$ de $\mathcal{V}_i$

⇒ méthode de Lloyd pour minimiser $J(\mathbf{X}_n) = \sum_{i=1}^n \int_{\mathbf{x} \in \mathcal{V}_i} \|\mathbf{x} - \mathbf{x}_i\|^2 d\mathbf{x}$ (Lloyd, 1982; Du et al., 1999)

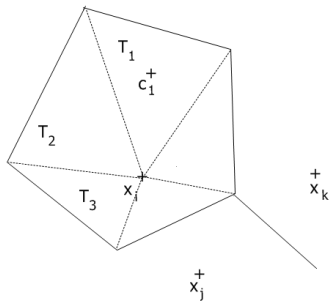
2) Remplacer $\mathbf{x}_i^k$ par $\mathbf{x}_i^{k+1}$ = centre $\mathbf{c}_E(\mathcal{V}_i)$ de l'ellipsoïde $\mathcal{E}_i^*$ de vol. min. contenant $\mathcal{V}_i$ ($\mathcal{E}_i^*$ = ellipsoïde de **John (1985)**, se calcule)

⇒ Ne converge probablement pas vers un min. local de $\Phi_{mM}(\mathbf{X}_n)$

3) Remplacer $\mathbf{x}_i^k$ par $\mathbf{x}_i^{k+1}$ = centre de gravité $\mathbf{c}_G(\mathcal{V}_i)$ de $\mathcal{V}_i$

⇒ méthode de Lloyd pour minimiser $J(\mathbf{X}_n) = \sum_{i=1}^n \int_{\mathbf{x} \in \mathcal{V}_i} \|\mathbf{x} - \mathbf{x}_i\|^2 d\mathbf{x}$ (**Lloyd, 1982; Du et al., 1999**)

Calcul centre de gravité $\mathbf{c}_G(\mathcal{V}_i)$: triangulation de Delaunay



$\mathcal{V}_i$ triangulé en $T_1, T_2, \dots$,

T_j a $d + 1$ sommets $\mathbf{s}_{j,k}$

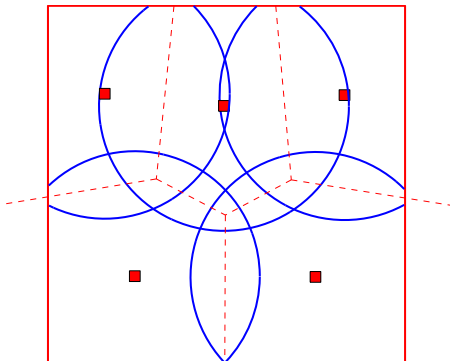
$$\mathbf{c}_G(\mathcal{V}_i) = \frac{\sum_j m_j \mathbf{c}_G(T_j)}{\sum_j m_j} \text{ avec}$$

$$m_j = \text{vol}(T_j) \text{ et } \mathbf{c}_G(T_j) = \frac{1}{d+1} \sum_{k=1}^{d+1} \mathbf{s}_{j,k}$$

- Possibilité d'utiliser autre norme, autre mesure qu'uniforme
- Possibilité de discrétiser $\mathcal{X}$ $\Rightarrow$ algorithme d'agrégation (*clustering*), `kmeans` par ex.

Ne converge pas vers un min. local de $\Phi_{mM}(\mathbf{X}_n)$, mais très simple à utiliser (peut servir à initialiser un autre algorithme)

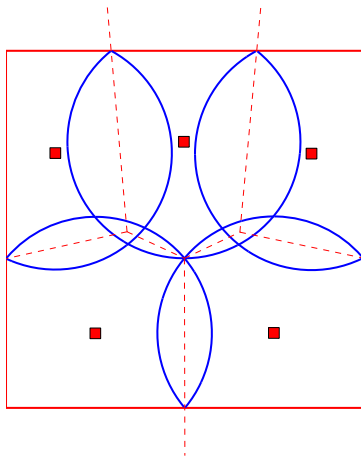
$d = 2$, $n = 5$, $\mathcal{X} = [0, 1]^2 \rightarrow \mathcal{X}_Q$, $Q = 1000$ (suite de Sobol'), `kmeans`



- Possibilité d'utiliser autre norme, autre mesure qu'uniforme
- Possibilité de discrétiser $\mathcal{X}$ $\Rightarrow$ algorithme d'agrégation (*clustering*), `kmeans` par ex.

Ne converge pas vers un min. local de $\Phi_{mM}(\mathbf{X}_n)$, mais très simple à utiliser (peut servir à initialiser un autre algorithme)

$d = 2$, $n = 5$, $\mathcal{X} = [0, 1]^2$, gradient généralisé: un peu mieux





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

Facile à calculer (à partir des distances d_{ij} entre paires de points)

$\Phi_{Mm}(\mathbf{X}_n) =$: minimum de fonctions convexes $\Rightarrow$ **non-concave, non-differentiable !**

- ▢ Utiliser un outil d'optimisation globale (ex., recuit simulé)
- ▢ **Descente locale** avec initialisation ad'hoc
(par ex. aléatoire $\Rightarrow$ *multistart*)



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

Facile à calculer (à partir des distances d_{ij} entre paires de points)

$\Phi_{Mm}(\mathbf{X}_n) = :$ minimum de fonctions convexes $\Rightarrow$ **non-concave, non-differentiable !**

▮ Utiliser un outil d'optimisation globale (ex., recuit simulé)

▮ **Descente locale** avec initialisation ad'hoc
(par ex. aléatoire $\Rightarrow$ *multistart*)

Problème assez difficile, mais :

$\Phi_{Mm}(\cdot)$ est globalement Lipschitz (avec constante 1) (Cortés and Bullo, 2005, 2009)

$\rightarrow \Phi_{Mm}(\cdot)$ est différentiable presque partout



Sous-différentiel $\partial\Phi_{Mm}(\mathbf{X}_n)$ facile à calculer:

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

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

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

Algorithme de sous-gradient pour maximiser $\Phi_{Mm}(\cdot)$:

$$\mathbf{X}_n^{k+1} = \text{Proj}_{\mathcal{X}} [\mathbf{X}_n^k + \gamma_k \partial \Phi_{Mm}(\mathbf{X}_n^k)]$$

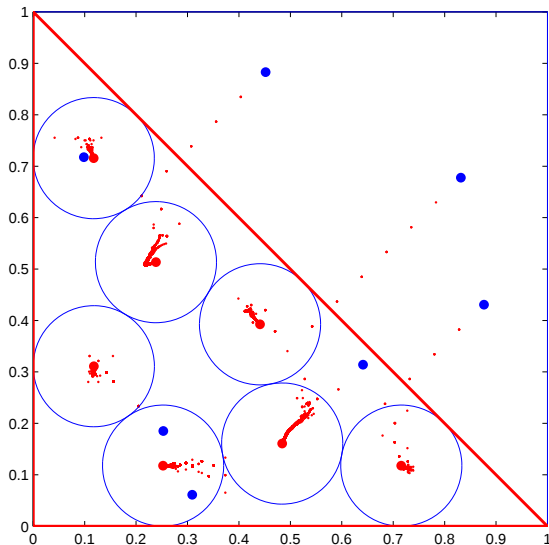
pour i, j tels que $\|\mathbf{x}_i - \mathbf{x}_j\| = \Phi_{Mm}(\mathbf{X}_n^k)$

et $\gamma_k \searrow 0$, $\sum_k \gamma_k = \infty$

On peut forcer les points à rester loin des bords de $\mathcal{X}$:

$$\Phi_{Mm \ B(\mathcal{X})}(\mathbf{X}_n) = \min\{\Phi_{Mm}(\mathbf{X}_n), 2 \min_i d[\mathbf{x}_i, \text{boundary}(\mathcal{X})]\}$$

Ex: maximisation locale de $\Phi_{Mm\ B(\mathcal{X})}(\mathbf{X}_n)$, $n = 7$, $\mathcal{X} = \text{simplex } 0 \leq x_1, 0 \leq x_2, x_1 + x_2 \leq 1$
 (rayon = $\frac{1}{2} \Phi_{Mm\ B(\mathcal{X})}(\mathbf{X}_n)$)



Quelques variations autour du sous-gradient:

1) (Cortés and Bullo, 2005, 2009) : évolution équa. diff. à temps continu (sans $\gamma_k \searrow 0$)

Quelques variations autour du sous-gradient:

1) (Cortés and Bullo, 2005, 2009) : évolution équa. diff. à temps continu (sans $\gamma_k \searrow 0$)

2) Méthode de plus grande pente: on utilise tout le sous-différentiel $\partial\Phi_{Mm}(\mathbf{X}_n)$

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

au lieu d'un seul sous-gradient $\partial\Phi_{Mm\ ij}(\mathbf{X}_n)$

Quelques variations autour du sous-gradient:

1) (Cortés and Bullo, 2005, 2009) : évolution équa. diff. à temps continu (sans $\gamma_k \searrow 0$)

2) Méthode de plus grande pente: on utilise tout le sous-différentiel $\partial\Phi_{Mm}(\mathbf{X}_n)$

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

au lieu d'un seul sous-gradient $\partial\Phi_{Mm\ ij}(\mathbf{X}_n)$

➤ direction dans laquelle $\Phi_{Mm}(\cdot)$ augmente (programmation quadratique, ou linéaire)

➤ méthode des approximations successives (Dem'yanov, 1968; Dem'yanov and Malozemov, 1974)

➡ **convergence monotone** vers un point stationnaire de $\Phi_{Mm}(\cdot)$ (max. local), un peu lourd...

Quelques alternatives:

1) Maillage des points $\mathbf{X}_n$ par triangulation de Delaunay, forces de répulsion ad'hoc le long des arêtes et sur les bords de $\mathcal{X}$ (Persson and Strang, 2004)

Quelques alternatives:

- 1) Maillage des points $\mathbf{X}_n$ par triangulation de Delaunay, forces de répulsion ad'hoc le long des arêtes et sur les bords de $\mathcal{X}$ (Persson and Strang, 2004)
- 2) Algorithme de billard (Lubachevsky and Stillinger, 1990; Lubachevsky, 1991)

Principe :

$\mathbf{X}_n \rightarrow n$ boules dans $\mathcal{X}$

vitesses initiales aléatoires

chocs élastiques entre les boules et contre les parois

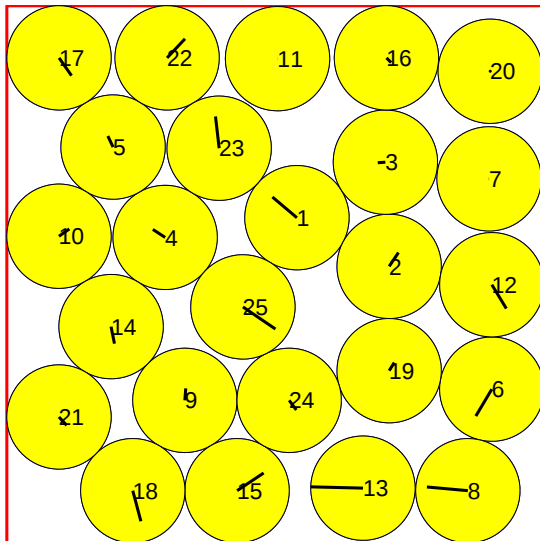
rayon $R(t)$ des boules croissant linéairement au cours du temps t

▮ blocage en un max. local de $\Phi_{Mm}(\cdot)$

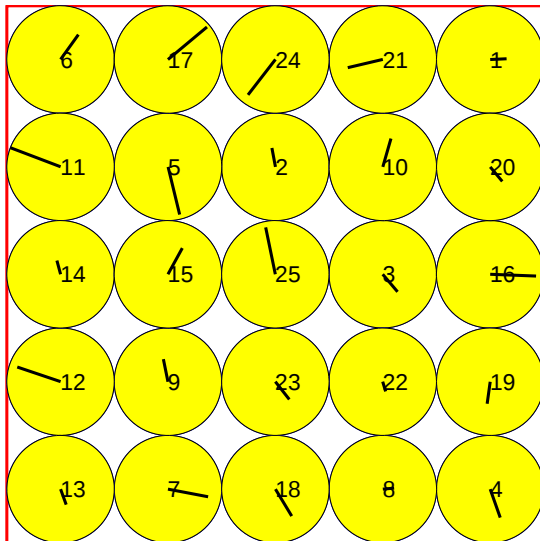
Assez performant pour $d = 2$ (si $R(t)$ ne croît pas trop vite. . .)

Ne semble pas très efficace pour $d > 2$

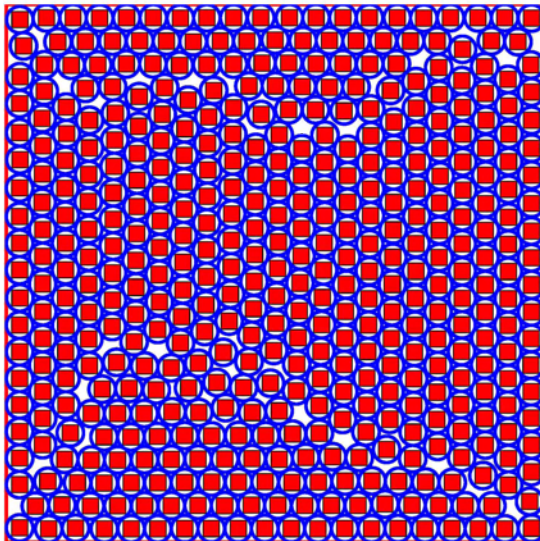
$n = 25$ ($R(t)$ croît trop vite)



$n = 25$ ($R(t)$ ne croît pas trop vite)



$$n = 441$$



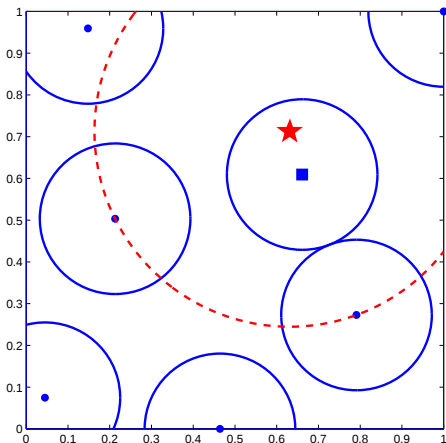
3) miniMax pour Maximin

Principe : répéter les étapes suivantes

- a) Choisir un point $\mathbf{x}_j$ de $\mathbf{X}_n$,
 trouver le point $\mathbf{x}^*$ de $\mathcal{X}$ tel que

$$\min_{i \neq j} \|\mathbf{x}^* - \mathbf{x}_i\| = \max_{\mathbf{x} \in \mathcal{X}} \min_{i \neq j} \|\mathbf{x} - \mathbf{x}_i\|$$

 ($\mathbf{x}^*$ le plus loin des $\mathbf{x}_i$ restants de $\mathbf{X}_n$,
 résultat du calcul de $\Phi_{mM}(\mathbf{X}_{n \setminus j})$)



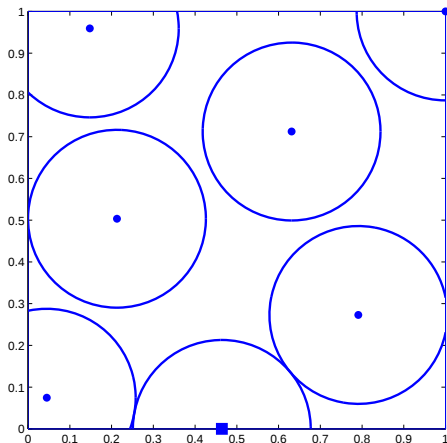
3) miniMax pour Maximin

Principe : répéter les étapes suivantes

b) Remplacer $\mathbf{x}_j$ par $\mathbf{x}^*$

— ce qui explique que

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



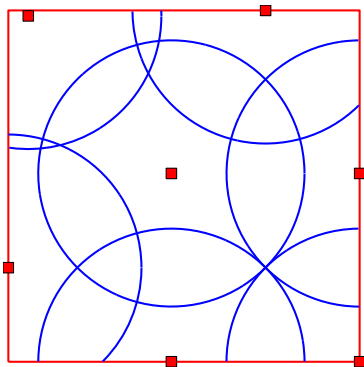
4) Maximin pour miniMax

Principe :

Les points d'un plan Maximin-optimal $\mathbf{X}_n^*$ ont tendance à être sur les bords de $\mathcal{X}$

▮ Appliquer une homothétie de centre $\mathbf{c} \in \text{int}(\mathcal{X})$ de rapport $1/(1 + \epsilon)$ aux point de $\mathbf{X}_n^*$ ($\mathcal{X} = [0, 1]^d$, $\mathbf{c} = \frac{1}{2} \mathbf{1}$ $\Rightarrow \mathbf{X}_n(\epsilon) = \mathbf{c} + \frac{1}{1+\epsilon} (\mathbf{X}_n^* - \mathbf{c})$)

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



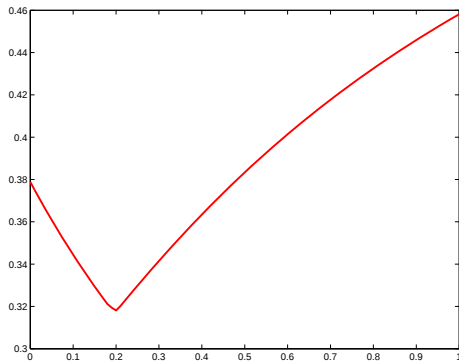
4) Maximin pour miniMax

Principe :

Les points d'un plan Maximin-optimal $\mathbf{X}_n^*$ ont tendance à être sur les bords de $\mathcal{X}$

⇒ Appliquer une homothétie de centre $\mathbf{c} \in \text{int}(\mathcal{X})$ de rapport $1/(1 + \epsilon)$ aux point de $\mathbf{X}_n^*$ ($\mathcal{X} = [0, 1]^d$, $\mathbf{c} = \frac{1}{2} \mathbf{1}$ ⇒ $\mathbf{X}_n(\epsilon) = \mathbf{c} + \frac{1}{1+\epsilon} (\mathbf{X}_n^* - \mathbf{c})$)

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



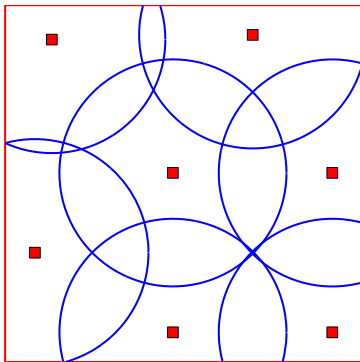
4) Maximin pour miniMax

Principe :

Les points d'un plan Maximin-optimal $\mathbf{X}_n^*$ ont tendance à être sur les bords de $\mathcal{X}$

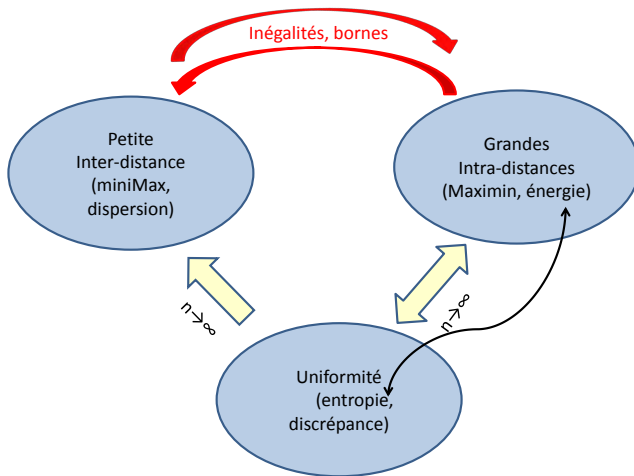
⇒ Appliquer une homothétie de centre $\mathbf{c} \in \text{int}(\mathcal{X})$ de rapport $1/(1 + \epsilon)$ aux point de $\mathbf{X}_n^*$ ($\mathcal{X} = [0, 1]^d$, $\mathbf{c} = \frac{1}{2} \mathbf{1}$ ⇒ $\mathbf{X}_n(\epsilon) = \mathbf{c} + \frac{1}{1+\epsilon} (\mathbf{X}_n^* - \mathbf{c})$)

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



►► Maximin régularisé

Relations entre Φ_{Mm} et Φ_{mM} ($d \geq 2$)



Relations entre Φ_{Mm} et Φ_{mM} ($d \geq 2$)

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

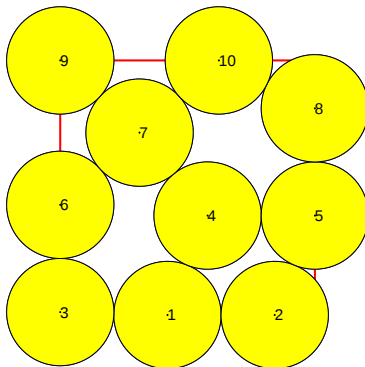
$$\triangleright \Phi_{mM}^* \triangleq \min_{\mathbf{X}} \Phi_{mM}(\mathbf{X}), \Phi_{Mm}^* \triangleq \max_{\mathbf{X}} \Phi_{Mm}(\mathbf{X}) \triangleleft$$

Relations entre Φ_{Mm} et Φ_{mM} ($d \geq 2$)

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

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

- $\frac{1}{2}\Phi_{Mm}(\mathbf{X}) \leq \Phi_{mM}(\mathbf{X})$ pour tout $\mathbf{X}$ (les n boules $\mathcal{B}(X_i, \frac{1}{2}\Phi_{Mm}(\mathbf{X}))$ ne recouvrent pas $\mathcal{X}$)



- Recouvrement par des sphères $\Rightarrow nV_d(\Phi_{mM}^*)^d > 1$

$$R^* < \Phi_{mM}^*$$

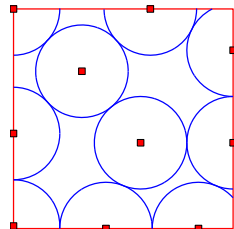
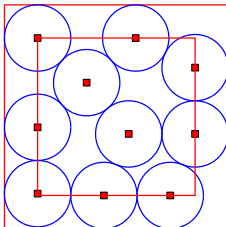
$$\text{avec } R^* = R^*(n) = (nV_d)^{-1/d}$$

- Recouvrement par des sphères $\Rightarrow nV_d(\Phi_{mM}^*)^d > 1$
- $\Phi_{mM}^* \leq \boxed{\Phi_{mM}(\mathbf{X}_{Mm}^*) \leq \Phi_{Mm}(\mathbf{X}_{Mm}^*)} = \Phi_{Mm}^*$
(preuve par contradiction)

$$R^* < \Phi_{mM}^* \leq \Phi_{Mm}^*$$

$$\text{avec } R^* = R^*(n) = (nV_d)^{-1/d}$$

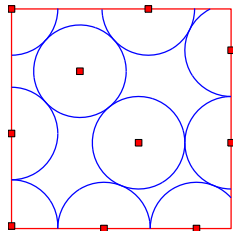
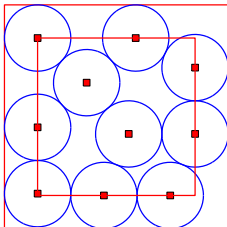
- Recouvrement par des sphères $\Rightarrow nV_d(\Phi_{mM}^*)^d > 1$
- $\Phi_{mM}^* \leq \boxed{\Phi_{mM}(\mathbf{X}_{Mm}^*) \leq \Phi_{Mm}(\mathbf{X}_{Mm}^*)} = \Phi_{Mm}^*$
(preuve par contradiction)
- Empilement de n sphères de rayon R dans $[0, 1]^d$
 $\Rightarrow nV_d R^d < 1$, i.e., $\boxed{R < R^*}$
et pour $n > \lceil 2^d/V_d \rceil$ ($R^* < \frac{1}{2}$)



$$\boxed{R^* < \Phi_{mM}^* \leq \Phi_{Mm}^* < \frac{2R^*}{1-2R^*}}$$

avec $R^* = R^*(n) = (nV_d)^{-1/d}$

- Recouvrement par des sphères $\Rightarrow nV_d(\Phi_{mM}^*)^d > 1$
- $\Phi_{mM}^* \leq \boxed{\Phi_{mM}(\mathbf{X}_{Mm}^*) \leq \Phi_{Mm}(\mathbf{X}_{Mm}^*)} = \Phi_{Mm}^*$
(preuve par contradiction)
- Empilement de n sphères de rayon R dans $[0, 1]^d$
 $\Rightarrow nV_d R^d < 1$, i.e., $\boxed{R < R^*}$
et pour $n > \lceil 2^d/V_d \rceil$ ($R^* < \frac{1}{2}$)



$$\boxed{R^* < \Phi_{mM}^* \leq \Phi_{Mm}^* < \frac{2R^*}{1-2R^*}}$$

avec $R^* = R^*(n) = (nV_d)^{-1/d}$

(et aussi $\Phi_{Mm}^* \leq \sqrt{d}$)

Remarques :

- De meilleures bornes existent pour $d = 2, 3$: empilement de n sphères de rayon R dans $[0, 1]^d \Rightarrow nV_d R^d < \delta_d = \text{packing density}$, avec
 - $\delta_2 = \pi\sqrt{3}/6 \simeq 0.9069$
[Lagrange, 1773 — pour structure en réseau, Tóth 1940 — résultat général]
 - $\delta_3 = \pi\sqrt{2}/6 \simeq 0.7405$
[Kepler 1611 — conjecture, Gauss 1831 — pour structure en réseau, Callister Hales 1998 — résultat général]

Remarques :

- De meilleures bornes existent pour $d = 2, 3$: empilement de n sphères de rayon R dans $[0, 1]^d \Rightarrow nV_d R^d < \delta_d = \text{packing density}$, avec
 - $\delta_2 = \pi\sqrt{3}/6 \simeq 0.9069$
[Lagrange, 1773 — pour structure en réseau, Tóth 1940 — résultat général]
 - $\delta_3 = \pi\sqrt{2}/6 \simeq 0.7405$
[Kepler 1611 — conjecture, Gauss 1831 — pour structure en réseau, Callister Hales 1998 — résultat général]
- Il existe des résultats récents (Wahl et al., 2014) sur la distribution de $\Phi_{Mm}(\mathbf{X}_n)$ quand les $\mathbf{x}_i$ sont i.i.d. uniformément dans $[0, 1]^d$

Remarques :

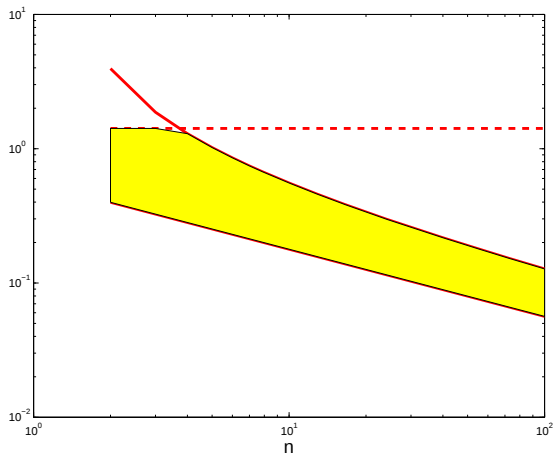
- De meilleures bornes existent pour $d = 2, 3$: empilement de n sphères de rayon R dans $[0, 1]^d \Rightarrow nV_d R^d < \delta_d = \text{packing density}$, avec
 - $\delta_2 = \pi\sqrt{3}/6 \simeq 0.9069$
[Lagrange, 1773 — pour structure en réseau, Tóth 1940 — résultat général]
 - $\delta_3 = \pi\sqrt{2}/6 \simeq 0.7405$
[Kepler 1611 — conjecture, Gauss 1831 — pour structure en réseau, Callister Hales 1998 — résultat général]
- Il existe des résultats récents (Wahl et al., 2014) sur la distribution de $\Phi_{mM}(\mathbf{X}_n)$ quand les $\mathbf{x}_i$ sont i.i.d. uniformément dans $[0, 1]^d$
- $R^*(n) < \Phi_{mM}^* \leq \Phi_{Mm}^* < \frac{2R^*(n)}{1-2R^*(n)} > 2R^*(n)$ ($R^*(n) = (nV_d)^{-1/d}$)
et le facteur 2 est très grand (!) : $R^*(n') = \frac{1}{2}R^*(n)$ pour $n' = 2^d n$
(si $\Phi_{mM}(\mathbf{X}_n) = \alpha$, alors en “copiant” 2^d fois $\mathbf{X}_n$, $\Phi_{mM}(\mathbf{X}_{2^d n}) = \frac{\alpha}{2}$)

► Maximin régularisé

Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 2$

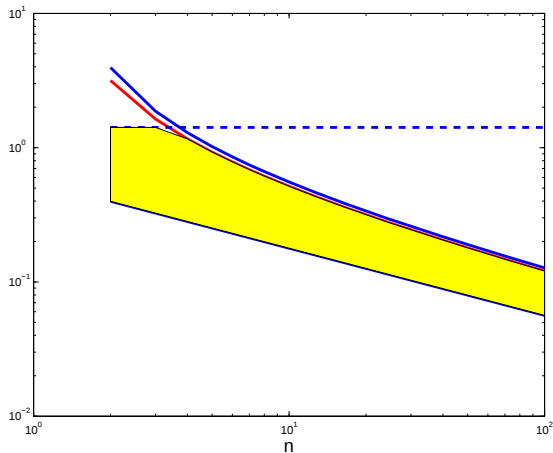
$d = 2$ utilisant

$$R^* < \Phi_{mM}^* \leq \Phi_{Mm}^* < \min\left\{\frac{2R^*}{1-2R^*}, \sqrt{d}\right\}$$



Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 2$

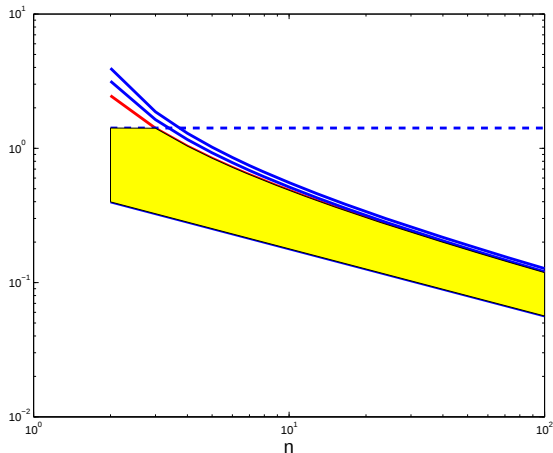
$d = 2$ avec de plus $\Phi_{Mm}^* < \frac{1}{3^{1/4} \sqrt{n/2-1}}$ (*packing density*)



Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 2$

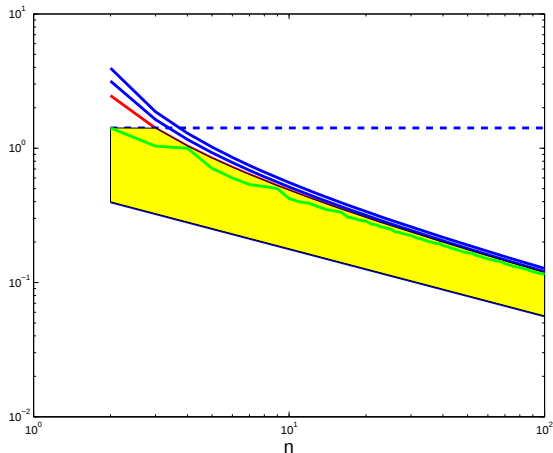
$d = 2$ avec de plus

$$\Phi_{Mm}^* < \frac{1 + \sqrt{1 + 2(n-1)/\sqrt{3}}}{n-1} \quad (\text{Oler, 1961})$$



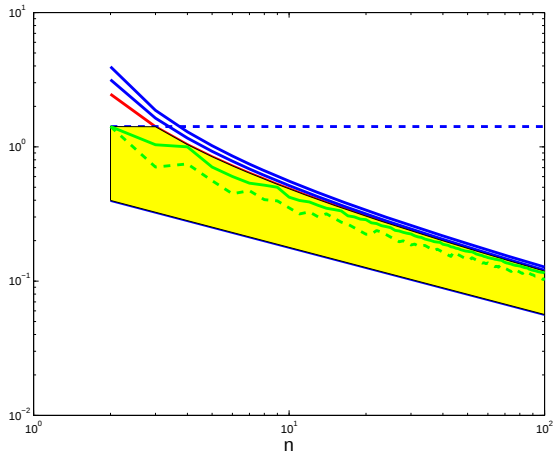
Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 2$

$d = 2$, en incluant Φ_{Mm}^* (prouvé jusqu'à $n = 30$ <http://www.packomania.com/>)



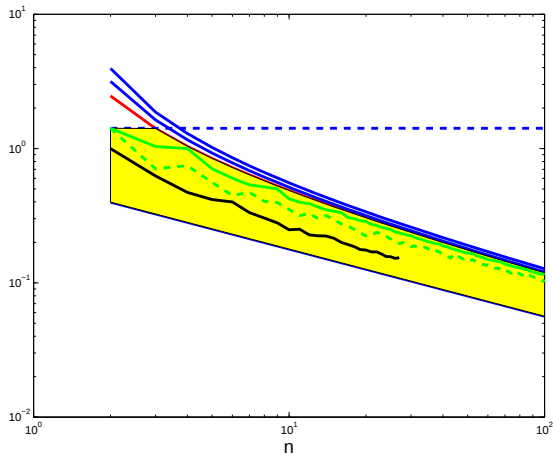
Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 2$

$d = 2$, en incluant Φ_{Mm-Lh}^* (prouvé jusqu'à $n = 70$ van Dam et al. (2007), <http://www.spacefillingdesigns.nl/>)



Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 2$

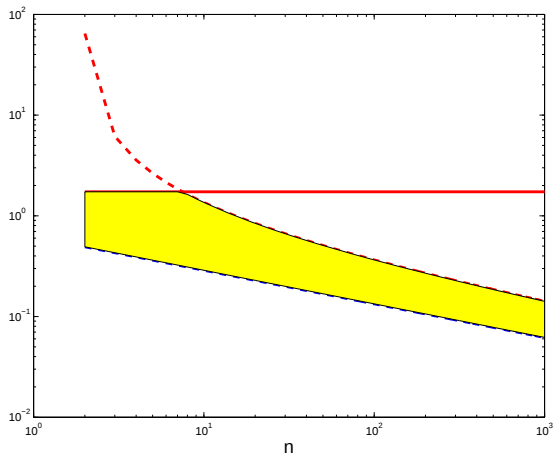
$d = 2$, en incluant Φ_{mM-Lh}^* (prouvé jusqu'à $n = 27$ (van Dam, 2008))



Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 3$

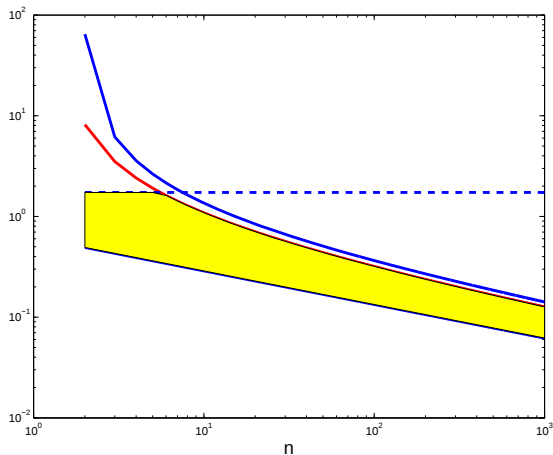
$d = 3$ utilisant

$$R^* < \Phi_{mM}^* \leq \Phi_{Mm}^* < \min\left\{\frac{2R^*}{1-2R^*}, \sqrt{d}\right\}$$



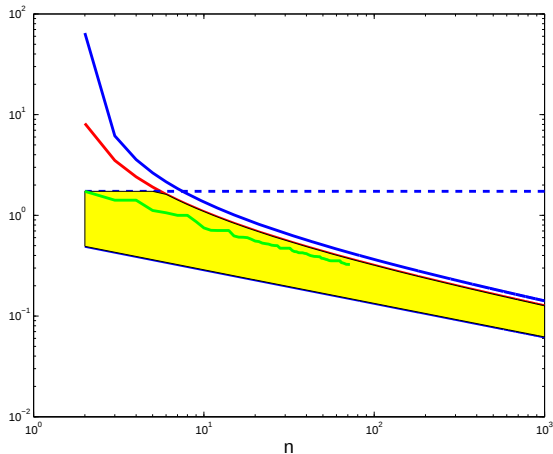
Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 3$

$d = 3$ avec de plus $\Phi_{Mm}^* < \frac{1}{(n/\sqrt{2})^{1/3}-1}$ (*packing density*)



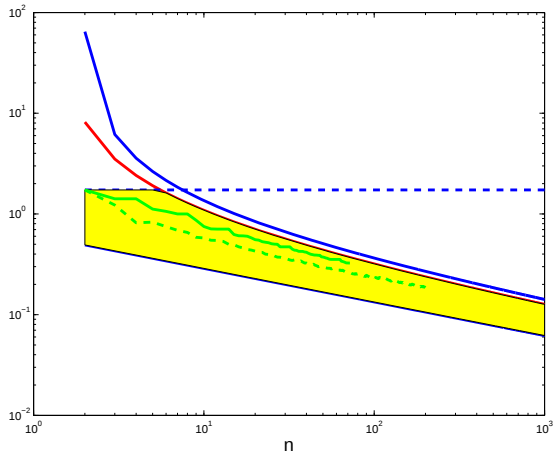
Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 3$

$d = 3$, incluant Φ_{Mm}^* (<http://www.randomwalk.de/sphere/incube/>)



Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 3$

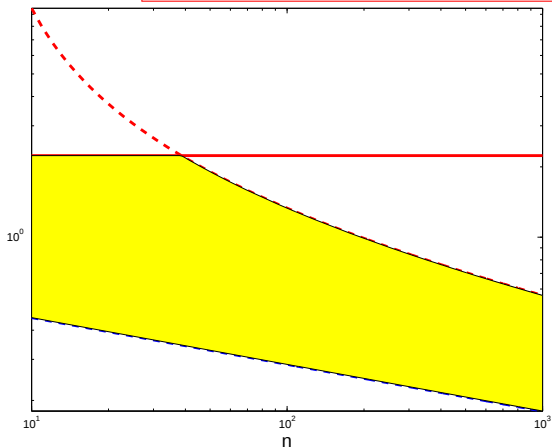
$d = 3$, incluant Φ_{Mm-Lh}^* (prouvé jusqu'à $n = 15$)
<http://www.spacefillingdesigns.nl/>



Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 5$

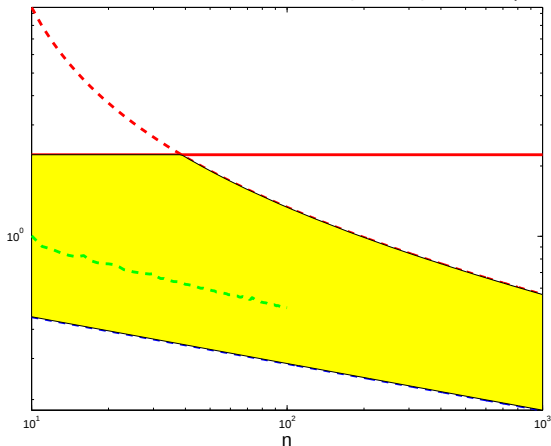
$d = 5$ utilisant

$$R^* < \Phi_{mM}^* \leq \Phi_{Mm}^* < \min\left\{\frac{2R^*}{1-2R^*}, \sqrt{d}\right\}$$



Bornes sur Φ_{Mm}^* et Φ_{mM}^* : $d = 5$

$d = 5$, incluant Φ_{Mm-Lh}^* (prouvé jusqu'à $n = 6$
<http://www.spacefillingdesigns.nl/>)



1.5 Maximin régularisé, énergie

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

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

avec $\mu_{ij} > 0$ et $\sum_{i < j} \mu_{ij} = 1$

1.5 Maximin régularisé, énergie

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

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

avec $\mu_{ij} > 0$ et $\sum_{i < j} \mu_{ij} = 1$

Alors $\underline{\Phi}_{[q]}(\mathbf{X}_n) \leq \Phi_{Mm}(\mathbf{X}_n) \leq \overline{\Phi}_{[q]}(\mathbf{X}_n) \leq \underline{\mu}^{-1/q} \underline{\Phi}_{[q]}(\mathbf{X}_n)$, $q > 0$, avec

$\underline{\mu} = \min_{i < j} \mu_{ij}$ (convergence monotone en q de chaque coté vers $\Phi_{Mm}(\mathbf{X})$ quand $q \rightarrow \infty$)

1.5 Maximin régularisé, énergie

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

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

avec $\mu_{ij} > 0$ et $\sum_{i < j} \mu_{ij} = 1$

Alors $\underline{\Phi}_{[q]}(\mathbf{X}_n) \leq \Phi_{Mm}(\mathbf{X}_n) \leq \overline{\Phi}_{[q]}(\mathbf{X}_n) \leq \underline{\mu}^{-1/q} \underline{\Phi}_{[q]}(\mathbf{X}_n)$, $q > 0$, avec

$\underline{\mu} = \min_{i < j} \mu_{ij}$ (convergence monotone en q de chaque coté vers $\Phi_{Mm}(\mathbf{X})$ quand $q \rightarrow \infty$)

Par continuité : $\overline{\Phi}_{[0]}(\mathbf{X}_n) = \exp \left[\sum_{i < j} \mu_{ij} \log(d_{ij}) \right]$

μ = mesure uniforme ($\mu_{ij} = \underline{\mu} = \binom{n}{2}^{-1}$ pour tout $i < j$) $\Rightarrow$

$$\frac{\Phi_{Mm}(\underline{\mathbf{X}}_{n[q]}^*)}{\Phi_{Mm}^*} \geq \left(\frac{n}{2}\right)^{-1/q},$$

avec $\underline{\mathbf{X}}_{n[q]}^*$ optimal pour $\underline{\Phi}_{[q]}$

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

μ = mesure uniforme ($\mu_{ij} = \underline{\mu} = \binom{n}{2}^{-1}$ pour tout $i < j$) $\Rightarrow$

$$\frac{\Phi_{Mm}(\underline{\mathbf{X}}_{n[q]}^*)}{\Phi_{Mm}^*} \geq \left(\frac{n}{2}\right)^{-1/q},$$

avec $\underline{\mathbf{X}}_{n[q]}^*$ optimal pour $\Phi_{[q]}$

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

$q = 2$: **critère d'énergie** de Audze and Eglais (1977)

pour $q \lesssim 5$ optimisation (plans Lh) plus facile que pour Φ_{Mm} (Morris and Mitchell, 1995)

👉 Utiliser le + petit q pour que les plans optimaux coïncident ?

μ = mesure uniforme ($\mu_{ij} = \underline{\mu} = \binom{n}{2}^{-1}$ pour tout $i < j$) $\Rightarrow$

$$\frac{\Phi_{Mm}(\underline{\mathbf{X}}_{n[q]}^*)}{\Phi_{Mm}^*} \geq \left(\frac{n}{2}\right)^{-1/q},$$

avec $\underline{\mathbf{X}}_{n[q]}^*$ optimal pour $\underline{\Phi}_{[q]}$

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

$q = 2$: **critère d'énergie** de Audze and Eglais (1977)

pour $q \lesssim 5$ optimisation (plans Lh) plus facile que pour Φ_{Mm} (Morris and Mitchell, 1995)

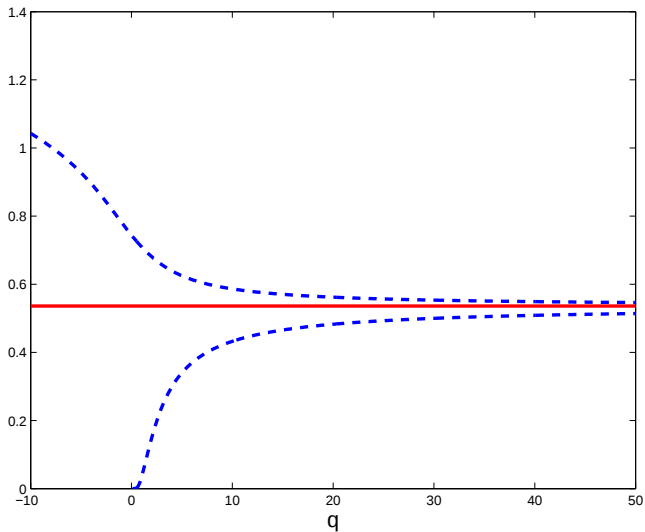
👉 Utiliser le + petit q pour que les plans optimaux coïncident ?

Version régularisée $\underline{\Phi}_{[q]}(\mathbf{X}_n)$: non-concave mais différentiable

➡ maximisation locale “facile” pour q pas trop grand ...

mais q doit être assez grand pour avoir une bonne approximation de $\Phi_{Mm}(\mathbf{X}_n)$!

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



6) Maximin régularisé avec Plus Proches Voisins (PPV)

On peut écrire $\Phi_{Mm}(\mathbf{X}_n) = \min_i d_i^*$, avec $d_i^* \triangleq \min_{j \neq i} \|\mathbf{x}_i - \mathbf{x}_j\|$
 = distance au PPV de X_i (*nearest neighbor*)
 (on ne régularise que le $\min_i$)

On définit

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

6) Maximin régularisé avec Plus Proches Voisins (PPV)

On peut écrire $\Phi_{Mm}(\mathbf{X}_n) = \min_i d_i^*$, avec $d_i^* \triangleq \min_{j \neq i} \|\mathbf{x}_i - \mathbf{x}_j\|$
 = distance au PPV de X_i (*nearest neighbor*)
 (on ne régularise que le $\min_i$)

On définit

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

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

(convergence monotone en q de chaque coté vers $\Phi_{Mm}(\mathbf{X}_n)$ quand $q \rightarrow \infty$)

Par continuité : $\bar{\Phi}_{[PPV,0]}(\mathbf{X}_n) = \exp \left[\sum_{i=1}^n \frac{\log(d_i^*)}{n} \right]$

Efficacité Maximin :

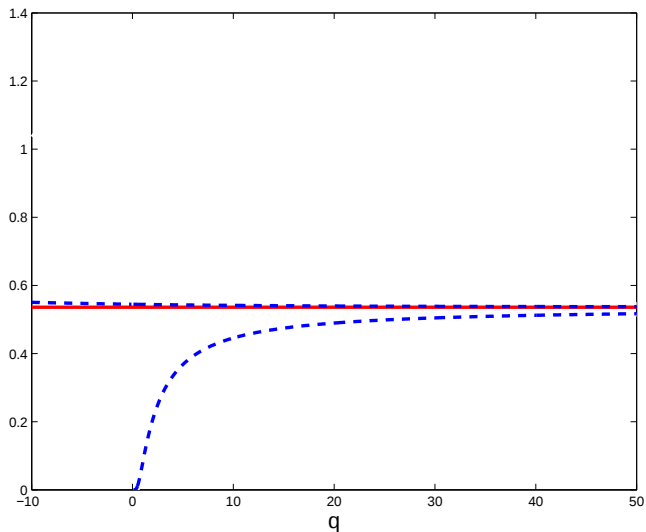
$$\frac{\Phi_{Mm}(\underline{\mathbf{X}}_{n[PPV,q]}^*)}{\Phi_{Mm}^*} \geq n^{-1/q},$$

avec $\underline{\mathbf{X}}_{n[PPV,q]}^*$ optimal pour $\Phi_{[PPV,q]}$

Efficacité Maximin $> 1 - \epsilon$ pour $q > \frac{\log(n)}{\epsilon}$
 $\rightarrow$ gain d'un facteur 2 par rapport à $\Phi_{[q]}(\mathbf{X}_n)$
 (on n'a régularisé qu'un seul min)

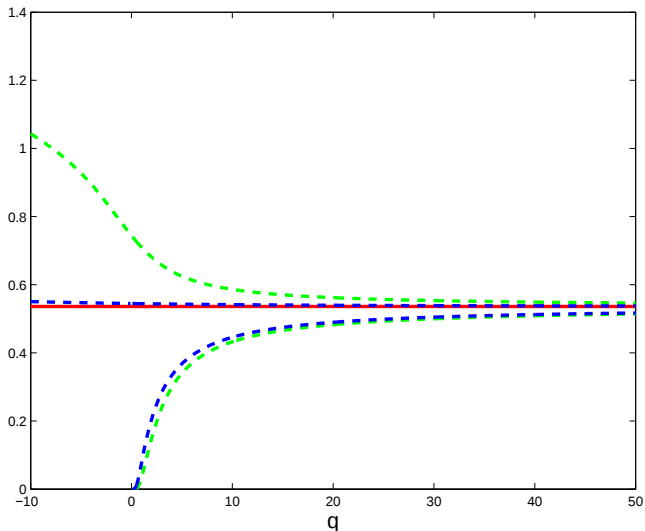
$\mathbf{X}_n$ Maximin-optimal, $n = 7$, $d = 2$:

Φ_{Mm} et régularisation par $\Phi_{[PPV,q]}$



$\mathbf{X}_n$ Maximin-optimal, $n = 7$, $d = 2$:

Φ_{Mm} et régularisation par $\underline{\Phi}_{[PPV,q]}$ et $\underline{\Phi}_{[q]}$



Remarque : mesures de régularité (et pas de remplissage !)

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

$R_1(\mathbf{X}_n) = 0$ pour une grille régulière

▲ invariant par changement d'échelle ▲

appelé “mesure de recouvrement”... mais ne mesure pas le recouvrement

Remarque : mesures de régularité (et pas de remplissage !)

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

$R_1(\mathbf{X}_n) = 0$ pour une grille régulière

▲ invariant par changement d'échelle ▲

appelé “mesure de recouvrement”... mais ne mesure pas le recouvrement

$$R_2(\mathbf{X}_n) \triangleq \frac{\max_{1 \leq i \leq n} d_i^*}{\min_{1 \leq i \leq n} d_i^*} (\geq 1)$$

$R_2(\mathbf{X}_n) = 1$ pour une grille régulière

▲ invariant par changement d'échelle ▲

2 Uniformité: quasi Monte-Carlo, discrédance

2.1 Entropie et graphes optimaux

Considérons $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ comme un échantillon de taille n de variables $\mathbf{x}_i$ i.i.d. dans $\mathcal{X}$ avec la densité de proba. $\varphi(\cdot)$

Entropie de Rényi de $\varphi(\cdot)$ d'ordre α :

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

Entropie de Tsallis de $\varphi(\cdot)$ d'ordre α :

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

et leur limite pour $\alpha \rightarrow 1$ donne

$$H_1(\varphi) \triangleq - \log \int \varphi(\mathbf{x}) \log[\varphi(\mathbf{x})] d\mathbf{x} \quad (\text{entropie de Shannon})$$

2 Uniformité: quasi Monte-Carlo, discrédance

2.1 Entropie et graphes optimaux

Considérons $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ comme un échantillon de taille n de variables $\mathbf{x}_i$ i.i.d. dans $\mathcal{X}$ avec la densité de proba. $\varphi(\cdot)$

Entropie de Rényi de $\varphi(\cdot)$ d'ordre α :

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

Entropie de Tsallis de $\varphi(\cdot)$ d'ordre α :

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

et leur limite pour $\alpha \rightarrow 1$ donne

$$H_1(\varphi) \triangleq - \log \int \varphi(\mathbf{x}) \log[\varphi(\mathbf{x})] d\mathbf{x} \quad (\text{entropie de Shannon})$$

► Pour $\alpha > 0$, $H_{\alpha}^*(\varphi)$, $H_{\alpha}(\varphi)$ maximales pour φ uniforme sur $\mathcal{X}$

2 Uniformité: quasi Monte-Carlo, discrédance

2.1 Entropie et graphes optimaux

Considérons $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ comme un échantillon de taille n de variables $\mathbf{x}_i$ i.i.d. dans $\mathcal{X}$ avec la densité de proba. $\varphi(\cdot)$

Entropie de Rényi de $\varphi(\cdot)$ d'ordre α :

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

Entropie de Tsallis de $\varphi(\cdot)$ d'ordre α :

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

et leur limite pour $\alpha \rightarrow 1$ donne

$$H_1(\varphi) \triangleq -\log \int \varphi(\mathbf{x}) \log[\varphi(\mathbf{x})] d\mathbf{x} \quad (\text{entropie de Shannon})$$

➤ Pour $\alpha > 0$, $H_{\alpha}^*(\varphi)$, $H_{\alpha}(\varphi)$ maximales pour φ uniforme sur $\mathcal{X}$

⇒ Construire un estimateur $\hat{H}_{n\alpha}$ de $H_{\alpha}^*(\varphi)$ à partir de $\mathbf{X}_n$,
puis prendre $\hat{H}_{n\alpha}$ comme critère à maximiser par rapport à $\mathbf{X}_n$

▲ Ne pas confondre ces crit res d'entropie (de la “distribution des x_i ”) avec le crit re d'entropie pour les processus Gaussiens, voir § II-1.3 (m me si des liens existent)

▲ Ne pas confondre ces crit res d'entropie (de la “distribution des $\mathbf{x}_i$ ”) avec le crit re d'entropie pour les processus Gaussiens, voir § II-1.3 (m me si des liens existent)

1) Id e assez naturelle : m thode de substitution (*plug-in*)

- construire un estimateur   noyau $\hat{\varphi}_n(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n K_{\sigma^2}(\mathbf{x} - \mathbf{x}_i)$,
 $K_{\sigma^2}(\cdot)$ = d.d.p. de moyenne 0 et variance σ^2 (suffisamment petite)
- Utiliser $H_\alpha(\hat{\varphi}_n)$, $\alpha > 0$, comme crit re   maximiser

Jourdan and Franco (2010) utilisent $H_1(\varphi)$, assez co teux pour d grand

▲ Ne pas confondre ces critères d'entropie (de la “distribution des $\mathbf{x}_i$ ”) avec le critère d'entropie pour les processus Gaussiens, voir § II-1.3 (même si des liens existent)

1) Idée assez naturelle : méthode de substitution (*plug-in*)

- construire un estimateur à noyau $\hat{\varphi}_n(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n K_{\sigma^2}(\mathbf{x} - \mathbf{x}_i)$,
 $K_{\sigma^2}(\cdot)$ = d.d.p. de moyenne 0 et variance σ^2 (suffisamment petite)
- Utiliser $H_\alpha(\hat{\varphi}_n)$, $\alpha > 0$, comme critère à maximiser

Jourdan and Franco (2010) utilisent $H_1(\varphi)$, assez coûteux pour d grand

Une particularité de H_2 : si $K_{\sigma^2}(\cdot)$ correspond à $\mathcal{N}(0, \sigma^2)$, alors

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

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

= critère intra-distances !

2) Graphes “optimaux” :

(Beardwood et al., 1959): $\mathbf{x}_i$ i.i.d. avec d.d.p. φ , graphe du Voyageur de Commerce $\mathcal{G}_{VC}(\mathbf{X}_n)$, ar tes e_i

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

Puis (Steele, 1981) pour d’autres fonctionnelles Euclidiennes sur $\mathbf{X}_n$, (Redmond and Yukich, 1994) utilisant la notion de quasi-additivit 

2) Graphes “optimaux” :

(Beardwood et al., 1959): $\mathbf{x}_i$ i.i.d. avec d.d.p. φ , graphe du Voyageur de Commerce $\mathcal{G}_{VC}(\mathbf{X}_n)$, ar tes e_i

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

Puis (Steele, 1981) pour d’autres fonctionnelles Euclidiennes sur $\mathbf{X}_n$, (Redmond and Yukich, 1994) utilisant la notion de quasi-additivit 

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

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

avec $\mathcal{G}(\mathbf{X}_n)$ arbre de longueur minimale (ALM, *Minimum Spanning Tree*), PPV, VC, Vorono , Delaunay, Sphere d’Influence, Gabriel... (diff rents types de convergence (L_p), diff rentes conditions sur φ et β ...)

En r sum : on construit un de ces graphes $\mathcal{G}$ sur $\mathbf{X}_n$, puis

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

En r sum : on construit un de ces graphes $\mathcal{G}$ sur $\mathbf{X}_n$, puis

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

► estimation de l'entropie de R nyi entropy $H_\alpha^*(\varphi)$ pour $\alpha = (d - \beta)/d$

$$1 \leq \beta < d \Rightarrow 0 < \alpha \leq 1 - 1/d$$

maximiser $H_\alpha \Leftrightarrow$ maximiser $\Phi_{\mathcal{G},\beta}(\mathbf{X}_n)$ (intra-distances)

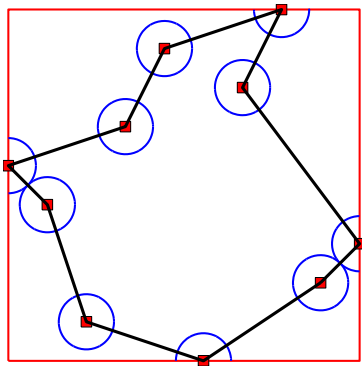
En r sum : on construit un de ces graphes $\mathcal{G}$ sur $\mathbf{X}_n$, puis

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

- estimation de l'entropie de R nyi entropy $H_\alpha^*(\varphi)$ pour $\alpha = (d - \beta)/d$
 $1 \leq \beta < d \Rightarrow 0 < \alpha \leq 1 - 1/d$

maximiser $H_\alpha \Leftrightarrow$ maximiser $\Phi_{\mathcal{G},\beta}(\mathbf{X}_n)$ (intra-distances)

$\mathbf{X}_n$ un Lh avec $n = 10$, $d = 2$: $\mathcal{G}_{VC}(\mathbf{X}_n)$



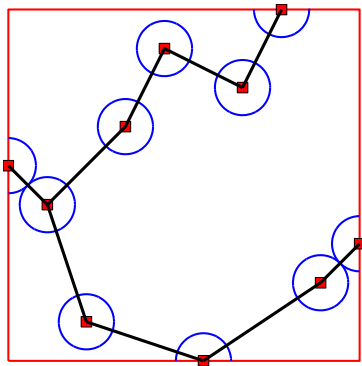
En r sum : on construit un de ces graphes $\mathcal{G}$ sur $\mathbf{X}_n$, puis

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

- estimation de l'entropie de R nyi entropy $H_\alpha^*(\varphi)$ pour $\alpha = (d - \beta)/d$
 $1 \leq \beta < d \Rightarrow 0 < \alpha \leq 1 - 1/d$

maximiser $H_\alpha \Leftrightarrow$ maximiser $\Phi_{\mathcal{G},\beta}(\mathbf{X}_n)$ (intra-distances)

$\mathbf{X}_n$ un Lh avec $n = 10$, $d = 2$: $\mathcal{G}_{ALM}(\mathbf{X}_n)$



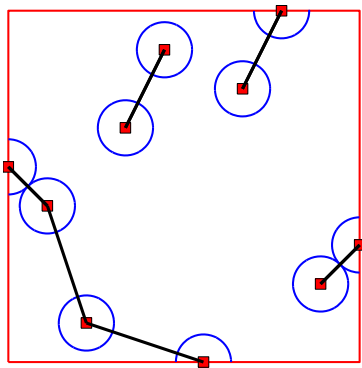
En r sum : on construit un de ces graphes $\mathcal{G}$ sur $\mathbf{X}_n$, puis

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

- estimation de l'entropie de R nyi $H_\alpha^*(\varphi)$ pour $\alpha = (d - \beta)/d$
 $1 \leq \beta < d \Rightarrow 0 < \alpha \leq 1 - 1/d$

maximiser $H_\alpha \Leftrightarrow$ maximiser $\Phi_{\mathcal{G},\beta}(\mathbf{X}_n)$ (intra-distances)

$\mathbf{X}_n$ un Lh avec $n = 10$, $d = 2$: $\mathcal{G}_{PPV}(\mathbf{X}_n)$



Graphe ALM : Franco (2008); Franco et al. (2009) utilisent une repr sentation dans le plan d fini par

la moyenne $E_n = (1/n) \sum_{e_i \in \mathcal{G}_{ALM}(\mathbf{x})} |e_i|$ et

l' cart type $S_n = (\text{var}_{\mathcal{G}_{ALM}(\mathbf{x})} \{|e_i|\})^{1/2}$

pour classifier diff rents types de plans $\mathbf{X}_n$ *space filling* (on pourrait utiliser un autre graphe $\mathcal{G}$)

Graphe ALM : Franco (2008); Franco et al. (2009) utilisent une repr sentation dans le plan d fini par

la moyenne $E_n = (1/n) \sum_{e_i \in \mathcal{G}_{ALM}(\mathbf{x})} |e_i|$ et

l' cart type $S_n = (\text{var}_{\mathcal{G}_{ALM}(\mathbf{x})} \{|e_i|\})^{1/2}$

pour classifier diff rents types de plans $\mathbf{X}_n$ *space filling* (on pourrait utiliser un autre graphe $\mathcal{G}$)

Graphe des PPV : $|e_i| = d_i^*$

maximiser $H_\alpha \Leftrightarrow$ maximiser $\overline{\Phi}_{[PPV,q]}(\mathbf{X}_n)$, $q = -\beta = d(\alpha - 1) \in (-d, -1]$

Graphe ALM : Franco (2008); Franco et al. (2009) utilisent une repr sentation dans le plan d fini par

la moyenne $E_n = (1/n) \sum_{e_i \in \mathcal{G}_{ALM}(\mathbf{x})} |e_i|$ et

l' cart type $S_n = (\text{var}_{\mathcal{G}_{ALM}(\mathbf{x})} \{|e_i|\})^{1/2}$

pour classifier diff rents types de plans $\mathbf{X}_n$ *space filling* (on pourrait utiliser un autre graphe $\mathcal{G}$)

Graphe des PPV : $|e_i| = d_i^*$

maximiser $H_\alpha \Leftrightarrow$ maximiser $\bar{\Phi}_{[PPV,q]}(\mathbf{X}_n)$, $q = -\beta = d(\alpha - 1) \in (-d, -1]$

(Leonenko et al., 2008) : $\frac{d}{n} \sum_{i=1}^n \log(d_i^*) \rightarrow H_1(\varphi)$, $n \rightarrow \infty$

maximiser $H_1 \Leftrightarrow$ maximiser $\bar{\Phi}_{[PPV,0]}(\mathbf{X}_n) = \exp \left[\sum_{i=1}^n \frac{\log(d_i^*)}{n} \right]$

maximiser $H_\alpha \Leftrightarrow$ maximiser $\bar{\Phi}_{[PPV,q]}(\mathbf{X}_n)$, $q = d(\alpha - 1) \in (-d, 1/2) \setminus \{0\}$

= maximiser un crit re intra-distances

▲ La situation est diff rente si on consid re tous les d_{ij} ▲

Si $q < d$, maximiser $\overline{\Phi}_{[q]}(\mathbf{X}_n)$ ne conduit pas   une distribution uniforme des $\mathbf{x}_i^*$
($n \rightarrow \infty$)

▲ La situation est différente si on considère tous les d_{ij} ▲

Si $q < d$, maximiser $\overline{\Phi}_{[q]}(\mathbf{X}_n)$ ne conduit pas à une distribution uniforme des $\mathbf{x}_i^*$ ($n \rightarrow \infty$)

Joseph et al. (2014) : $\mathbf{X}_n^*$ maximise $\underbrace{\Phi_{Mm}(\mathbf{X}_n)}_{=\overline{\Phi}_{[\infty]}(\mathbf{X}_n)} \sim_{n \rightarrow \infty} \text{uniforme}$

Version continue : q -énergie d'une distribution ξ

$$E_q(\xi) \triangleq \int_{\mathcal{X}^2} \|\mathbf{x} - \mathbf{y}\|^{-q} \xi(d\mathbf{x}) \xi(d\mathbf{y}), \quad q > 0$$

$$\text{et } \tilde{\Phi}_{[q]}(\xi) \triangleq [E_q(\xi)]^{-1/q} = \text{version continue de } \overline{\Phi}_{[q]}(\cdot)$$

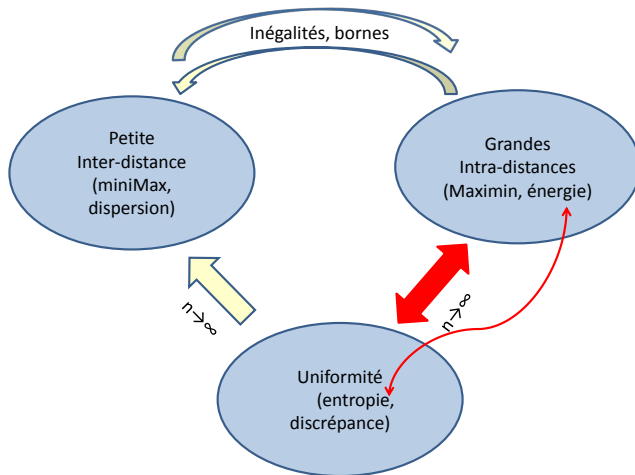
$$E_0(\xi) \triangleq - \int_{\mathcal{X}^2} \log[\|\mathbf{x} - \mathbf{y}\|] \xi(d\mathbf{x}) \xi(d\mathbf{y})$$

$$\text{et } \tilde{\Phi}_{[0]}(\xi) \triangleq \exp[-E_q(\xi)] = \text{version continue de } \overline{\Phi}_{[0]}(\cdot)$$

Alors (Hardin and Saff, 2004) :

$$E_q(\xi) \text{ minimum} \Leftrightarrow \tilde{\Phi}_{[q]}(\xi) \text{ maximum} \quad \begin{cases} \text{pour } \xi^* \text{ uniforme si } q \geq d \\ \text{pour } \xi^* = \text{mesure de } q\text{-équilibre} \\ \text{(non uniforme) si } 0 \leq q < d \end{cases}$$

- ⇒ Pour $\Phi_{Mm}(\cdot)$, ou $\bar{\Phi}_{[q]}(\cdot)$ avec $q > d$, ou $\bar{\Phi}_{[PPV,q]}(\cdot)$ avec $q > -d$,
la distribution de $\mathbf{X}_n$ doit être proche de l'uniforme



2.2 Discrédance : motivation (suivant le plan de (Niederreiter, 1992, Chap. 1,2))

— véritable monument comportant 371 références

[▶ QMC](#)

2.2 Discrédance : motivation (suivant le plan de (Niederreiter, 1992, Chap. 1,2) — véritable monument comportant 371 références) [▶ QMC](#)

Intégration par la méthode des trapèzes

- $d = 1, \mathbb{I}_1 = [0, 1]$

$$\int_0^1 f(u) \, du \simeq \sum_{i=0}^n w_i f(i/n), \quad w_0 = w_n = 1/n, \quad w_i = 1/(2n), \quad 1 \leq i \leq n-1$$

↪ erreur = $\mathcal{O}(n^{-2})$ quand f de classe C_2

2.2 Discrédance : motivation (suivant le plan de (Niederreiter, 1992, Chap. 1,2) — véritable monument comportant 371 références) ► QMC

Intégration par la méthode des trapèzes

- $d = 1$, $\mathbb{I}_1 = [0, 1]$

$$\int_0^1 f(u) \, du \simeq \sum_{i=0}^n w_i f(i/n), \quad w_0 = w_n = 1/n, \quad w_i = 1/(2n), \quad 1 \leq i \leq n-1$$

$\hookrightarrow$ erreur = $\mathcal{O}(n^{-2})$ quand f de classe C_2

- dimension $d > 1$, $\mathbb{I}_d \triangleq [0, 1]^d$

$$\int_{\mathbb{I}_d} f(\mathbf{u}) \, d\mathbf{u} \simeq \sum_{i_1=0}^m \cdots \sum_{i_d=0}^m w_{i_1} \cdots w_{i_d} f(i_1/m, \dots, i_d/m)$$

$\hookrightarrow$ erreur = $\mathcal{O}(m^{-2})$ quand $\partial^2 f / \partial u_\ell^2$ continues, $\ell = 1, \dots, d$

mais... $n = (m+1)^d$ points d'évaluations ("observations")

$$\Rightarrow \text{erreur} = \mathcal{O}(n^{-2/d})$$

2.2 Discrédance : motivation (suivant le plan de (Niederreiter, 1992, Chap. 1,2) — véritable monument comportant 371 références) ▶ QMC

Intégration par la méthode des trapèzes

- $d = 1$, $\mathbb{I}_1 = [0, 1]$

$$\int_0^1 f(u) \, du \simeq \sum_{i=0}^n w_i f(i/n), \quad w_0 = w_n = 1/n, \quad w_i = 1/(2n), \quad 1 \leq i \leq n-1$$

$\hookrightarrow$ erreur = $\mathcal{O}(n^{-2})$ quand f de classe C_2

- dimension $d > 1$, $\mathbb{I}_d \triangleq [0, 1]^d$

$$\int_{\mathbb{I}_d} f(\mathbf{u}) \, d\mathbf{u} \simeq \sum_{i_1=0}^m \cdots \sum_{i_d=0}^m w_{i_1} \cdots w_{i_d} f(i_1/m, \dots, i_d/m)$$

$\hookrightarrow$ erreur = $\mathcal{O}(m^{-2})$ quand $\partial^2 f / \partial u_\ell^2$ continues, $\ell = 1, \dots, d$

mais... $n = (m+1)^d$ points d'évaluations ("observations")

$$\Rightarrow \text{erreur} = \mathcal{O}(n^{-2/d})$$

⇒ Fléau de la dimension (*curse of dimensionality*):

$$\text{erreur} = \mathcal{O}(10^{-2}) \Rightarrow n = \mathcal{O}(10^d) !$$

Int gration par la m thode de Monte Carlo (MC)

$\mathcal{X}$ compact $\subset \mathbb{R}^d$, λ = mesure de Lebesgue ($\lambda(\mathcal{X})$ = volume de $\mathcal{X}$)

$$\begin{aligned}\int_{\mathcal{X}} f(\mathbf{u}) \, d\mathbf{u} &= \lambda(\mathcal{X}) \int_{\mathcal{X}} f(\mathbf{u}) \frac{d\mathbf{u}}{\lambda(\mathcal{X})} = \lambda(\mathcal{X}) \int_{\mathcal{X}} f \, d\mu \\ &\simeq \lambda(\mathcal{X}) \frac{1}{n} \sum_{i=1}^n f(x_i)\end{aligned}$$

avec x_i i.i.d. selon μ uniforme sur $\mathcal{X}$

Int gration par la m thode de Monte Carlo (MC)

$\mathcal{X}$ compact $\subset \mathbb{R}^d$, λ = mesure de Lebesgue ($\lambda(\mathcal{X})$ = volume de $\mathcal{X}$)

$$\begin{aligned} \int_{\mathcal{X}} f(\mathbf{u}) \, d\mathbf{u} &= \lambda(\mathcal{X}) \int_{\mathcal{X}} f(\mathbf{u}) \frac{d\mathbf{u}}{\lambda(\mathcal{X})} = \lambda(\mathcal{X}) \int_{\mathcal{X}} f \, d\mu \\ &\simeq \lambda(\mathcal{X}) \frac{1}{n} \sum_{i=1}^n f(x_i) \end{aligned}$$

avec x_i i.i.d. selon μ uniforme sur $\mathcal{X}$

$$\triangleright I(f) \triangleq \int_{\mathcal{X}} f \, d\mu, \hat{I}_n \triangleq \frac{1}{n} \sum_{i=1}^n f(x_i) \triangleleft$$

$$\Rightarrow E(\hat{I}_n) = I(f); \hat{I}_n \xrightarrow{\text{a.s.}} I(f), n \rightarrow \infty$$

$$f \in L_2(\mu) \quad (\sigma^2(f) = \int_{\mathcal{X}} (f - I(f))^2 \, d\mu < \infty)$$

$$\Rightarrow E([\hat{I}_n - I(f)]^2) = \frac{\sigma^2(f)}{n}$$

$$\hookrightarrow \text{erreur} \simeq \sigma(f) n^{-1/2}$$

Intégration par la méthode de Monte Carlo (MC)

$\mathcal{X}$ compact $\subset \mathbb{R}^d$, λ = mesure de Lebesgue ($\lambda(\mathcal{X})$ = volume de $\mathcal{X}$)

$$\begin{aligned} \int_{\mathcal{X}} f(\mathbf{u}) \, d\mathbf{u} &= \lambda(\mathcal{X}) \int_{\mathcal{X}} f(\mathbf{u}) \frac{d\mathbf{u}}{\lambda(\mathcal{X})} = \lambda(\mathcal{X}) \int_{\mathcal{X}} f \, d\mu \\ &\simeq \lambda(\mathcal{X}) \frac{1}{n} \sum_{i=1}^n f(x_i) \end{aligned}$$

avec x_i i.i.d. selon μ uniforme sur $\mathcal{X}$

$$\triangleright I(f) \triangleq \int_{\mathcal{X}} f \, d\mu, \hat{I}_n \triangleq \frac{1}{n} \sum_{i=1}^n f(x_i) \triangleleft$$

$$\Rightarrow E(\hat{I}_n) = I(f); \hat{I}_n \xrightarrow{\text{a.s.}} I(f), n \rightarrow \infty$$

$$f \in L_2(\mu) \quad (\sigma^2(f) = \int_{\mathcal{X}} (f - I(f))^2 \, d\mu < \infty)$$

$$\Rightarrow E([\hat{I}_n - I(f)]^2) = \frac{\sigma^2(f)}{n}$$

$$\hookrightarrow \text{erreur} \simeq \sigma(f) n^{-1/2}$$

➡ préférable à la méthode des trapèzes pour $d \geq 5$ (sans hypothèse de régularité sur f)

M thodes de Quasi-Monte Carlo (QMC)

 valuer f en des $\mathbf{x}_i$ d terministes dans $\mathcal{X} = \mathbb{I}_d \triangleq [0, 1]^d$:

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

pour toute fonction f int grable au sens de Riemann si les $\mathbf{x}_1, \mathbf{x}_2, \dots$ sont uniform ment distribu s dans $\mathbb{I}_d$

   il faut $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathcal{I}_{\mathbb{J}}(\mathbf{x}) = \lambda(\mathbb{J})$ pour tout sous-intervalle $\mathbb{J}$ de $\mathbb{I}_d$

Vitesse de convergence de $\hat{I}_n \rightarrow I(f)$?

M thodes de Quasi-Monte Carlo (QMC)

 valuer f en des $\mathbf{x}_i$ d terministes dans $\mathcal{X} = \mathbb{I}_d \triangleq [0, 1]^d$:

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

pour toute fonction f int grable au sens de Riemann si les $\mathbf{x}_1, \mathbf{x}_2, \dots$ sont uniform ment distribu s dans $\mathbb{I}_d$

   il faut $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathcal{I}_{\mathbb{J}}(\mathbf{x}) = \lambda(\mathbb{J})$ pour tout sous-intervalle $\mathbb{J}$ de $\mathbb{I}_d$

Vitesse de convergence de $\hat{I}_n \rightarrow I(f)$?

Pour tout $\mathbb{B} \subset \mathbb{I}_d$, $\mathbf{X}_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset \mathbb{I}_d^n$:

$$\triangleright A(\mathbb{B}, \mathbf{X}_n) \triangleq \sum_{i=1}^n \mathcal{I}_{\mathbb{B}}(\mathbf{x}_i) \triangleq \text{nb. de } \mathbf{x}_i \text{ dans } \mathbb{B} \triangleleft$$

M thodes de Quasi-Monte Carlo (QMC)

 valuer f en des $\mathbf{x}_i$ d terministes dans $\mathcal{X} = \mathbb{I}_d \triangleq [0, 1]^d$:

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

pour toute fonction f int grable au sens de Riemann si les $\mathbf{x}_1, \mathbf{x}_2, \dots$ sont uniform ment distribu s dans $\mathbb{I}_d$

  il faut $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathcal{I}_{\mathbb{J}}(\mathbf{x}) = \lambda(\mathbb{J})$ pour tout sous-intervalle $\mathbb{J}$ de $\mathbb{I}_d$

Vitesse de convergence de $\widehat{I}_n \rightarrow I(f)$?

Pour tout $\mathbb{B} \subset \mathbb{I}_d$, $\mathbf{X}_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset \mathbb{I}_d^n$:

$$\triangleright A(\mathbb{B}, \mathbf{X}_n) \triangleq \sum_{i=1}^n \mathcal{I}_{\mathbb{B}}(\mathbf{x}_i) \triangleq \text{nb. de } \mathbf{x}_i \text{ dans } \mathbb{B} \triangleleft$$

Discr pance (*Discrepancy*)

$$D_n(\mathcal{B}, \mathbf{X}_n) \triangleq \sup_{\mathbb{B} \in \mathcal{B}} \left| \frac{1}{n} A(\mathbb{B}, \mathbf{X}_n) - \lambda(\mathbb{B}) \right|$$

avec $\mathcal{B}$ une famille de sous ensembles de $\mathbb{I}_d$ Lebesgue-mesurables
(et $0 \leq D_n(\mathcal{B}, \mathbf{X}_n) \leq 1$)

Deux cas particuliers importants:

- **Discrédance à l'origine** (*star-discrepancy*) $D_n^*(\mathbf{X}_n) = D_n(\mathcal{B}, \mathbf{X}_n)$ quand $\mathcal{B}$ contient tous les sous-intervalles de la forme $\prod_{\ell=1}^d [0, u_\ell)$

$$D_n^*(\mathbf{X}_n) = \sup_{\mathbf{u} \in [0,1]^d} |F_n(\mathbf{u}) - F_U(\mathbf{u})|$$

($F_n(\cdot) \triangleq$ f. répartition empirique, $F_U(\mathbf{x}) \triangleq \prod_{\ell=1}^d \{\mathbf{x}\}_\ell =$ uniforme)

- **Discrédance extrême** (*extreme-discrepancy*) $D_n(\mathbf{X}_n) = D_n(\mathcal{B}, \mathbf{X}_n)$ quand $\mathcal{B}$ contient tous les sous-intervalles de la forme $\prod_{\ell=1}^d [u_\ell, v_\ell)$

Pour tout $\mathbf{X}_n \subset \mathbb{I}_d^n$, on a: $D_n^*(\mathbf{X}_n) \leq D_n(\mathbf{X}_n) \leq 2^d D_n^*(\mathbf{X}_n)$

Deux cas particuliers importants:

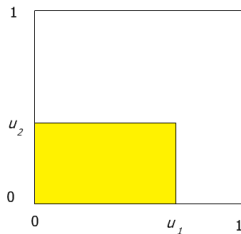
- **Discrédance à l'origine** (*star-discrepancy*) $D_n^*(\mathbf{X}_n) = D_n(\mathcal{B}, \mathbf{X}_n)$ quand $\mathcal{B}$ contient tous les sous-intervalles de la forme $\prod_{\ell=1}^d [0, u_\ell]$

$$D_n^*(\mathbf{X}_n) = \sup_{\mathbf{u} \in [0,1]^d} |F_n(\mathbf{u}) - F_U(\mathbf{u})|$$

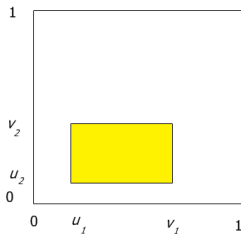
($F_n(\cdot) \triangleq$ f. répartition empirique, $F_U(\mathbf{x}) \triangleq \prod_{\ell=1}^d \{\mathbf{x}\}_\ell =$ uniforme)

- **Discrédance extrême** (*extreme-discrepancy*) $D_n(\mathbf{X}_n) = D_n(\mathcal{B}, \mathbf{X}_n)$ quand $\mathcal{B}$ contient tous les sous-intervalles de la forme $\prod_{\ell=1}^d [u_\ell, v_\ell]$

Pour tout $\mathbf{X}_n \subset \mathbb{I}_d^n$, on a: $D_n^*(\mathbf{X}_n) \leq D_n(\mathbf{X}_n) \leq 2^d D_n^*(\mathbf{X}_n)$



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



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

En dimension $d = 1$, avec $0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq 1$:

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

($\Rightarrow$ test d'uniformité de Kolmogorov-Smirnov)

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

et $D_n^*(\mathbf{X}_n) \geq \frac{1}{2n}$, $D_n(\mathbf{X}_n) \geq \frac{1}{n}$, avec égalité pour $\mathbf{X}_n^*$ tel que $x_i = \frac{2i-1}{2n} \forall i$

En dimension $d = 1$, avec $0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq 1$:

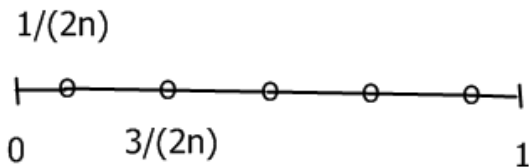
$$D_n^*(\mathbf{X}_n) = \frac{1}{2n} + \max_{1 \leq i \leq n} \left| x_i - \frac{2i-1}{2n} \right|$$

($\Rightarrow$ test d'uniformité de Kolmogorov-Smirnov)

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

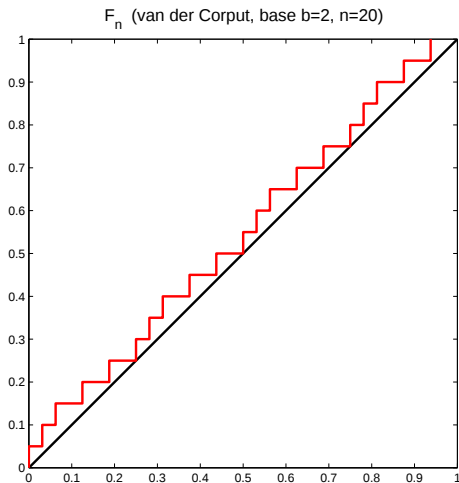
et $D_n^*(\mathbf{X}_n) \geq \frac{1}{2n}$, $D_n(\mathbf{X}_n) \geq \frac{1}{n}$, avec égalité pour $\mathbf{X}_n^*$ tel que $x_i = \frac{2i-1}{2n} \forall i$

$n=5$



fonct. répartition empirique $F_n(x)$ et $F_U(x)$

$d = 1$, $\mathbf{X}_{20} = 20$ premiers points d'une suite de van der Corput en base 2
(voir § 2.4)



Pourquoi est-ce important?

Pour $d = 1$, supposons f à variation bornée sur $[0, 1]$ ($V(f) \triangleq \int_0^1 |df(u)| < \infty$)

1) intégration par parties: $\int_0^1 f(u) du = f(1) - \int_0^1 u df(u)$

2) posons $0 = x_0 \leq x_1 \leq \dots \leq x_n \leq x_{n+1} = 1$,

alors $\frac{1}{n} \sum_{i=1}^n f(x_i) = f(1) - \sum_{i=0}^n \frac{i}{n} [f(x_{i+1}) - f(x_i)]$

3) pour tout $u \in [x_i, x_{i+1}]$, on a $|u - \frac{i}{n}| \leq D_n^*(\mathbf{X}_n)$

Pourquoi est-ce important?

Pour $d = 1$, supposons f à variation bornée sur $[0, 1]$ ($V(f) \triangleq \int_0^1 |df(u)| < \infty$)

1) intégration par parties: $\int_0^1 f(u) du = f(1) - \int_0^1 u df(u)$

2) posons $0 = x_0 \leq x_1 \leq \dots \leq x_n \leq x_{n+1} = 1$,

alors $\frac{1}{n} \sum_{i=1}^n f(x_i) = f(1) - \sum_{i=0}^n \frac{i}{n} [f(x_{i+1}) - f(x_i)]$

3) pour tout $u \in [x_i, x_{i+1}]$, on a $|u - \frac{i}{n}| \leq D_n^*(\mathbf{X}_n)$

$$\begin{aligned} \left| \int_0^1 f(u) du - \frac{1}{n} \sum_{i=1}^n f(x_i) \right| &\leq \left| \sum_{i=0}^n \left\{ \int_{x_i}^{x_{i+1}} u df(u) - \frac{i}{n} [f(x_{i+1}) - f(x_i)] \right\} \right| \\ &= \left| \sum_{i=0}^n \int_{x_i}^{x_{i+1}} \left(u - \frac{i}{n} \right) df(u) \right| \\ &\leq D_n^*(\mathbf{X}_n) V(f) \end{aligned}$$

⇒ Inégalité de Koksma (1942/1943) (non améliorable)

Par cons quent, pour $d = 1$

$$\left| \int_0^1 f(u) \, du - \frac{1}{n} \sum_{i=1}^n f(x_i) \right| \leq \frac{V(f)}{2n} \text{ pour } \mathbf{X}_n^* \text{ tel que } x_i = \frac{2i-1}{2n} \, \forall i$$

tandis que erreur MC $\simeq \frac{\sigma(f)}{n^{1/2}}$

Par cons quent, pour $d = 1$

$$\left| \int_0^1 f(u) du - \frac{1}{n} \sum_{i=1}^n f(x_i) \right| \leq \frac{V(f)}{2n} \text{ pour } \mathbf{X}_n^* \text{ tel que } x_i = \frac{2i-1}{2n} \forall i$$

tandis que erreur MC $\simeq \frac{\sigma(f)}{n^{1/2}}$

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

(in galit  de **Koksma-Hlawka (1961)**, non am liorable)

avec $V(f)$ = variation au sens de Hardy et Krause, et

$$\left| \int_{\mathbb{I}_d} f(\mathbf{u}) d\mathbf{u} - \frac{1}{n} \sum_{i=1}^n f(\mathbf{x}_i) \right| \leq \begin{cases} V(f) C_d \frac{(\log n)^{d-1}}{n} \\ \text{pour } \mathbf{X}_n = \text{ensemble de Hammersley} \\ V(f) C'_d \frac{(\log n)^d}{n} \\ \text{pour } \mathbf{X}_n = n \text{ premiers points} \\ \text{d'une suite de Halton (par ex.) } \mathbf{X}_\infty \end{cases}$$

voir § 2.4

Par cons quent, pour $d = 1$

$$\left| \int_0^1 f(u) du - \frac{1}{n} \sum_{i=1}^n f(x_i) \right| \leq \frac{V(f)}{2n} \text{ pour } \mathbf{X}_n^* \text{ tel que } x_i = \frac{2i-1}{2n} \forall i$$

tandis que erreur MC $\simeq \frac{\sigma(f)}{n^{1/2}}$

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

(in galit  de **Koksma-Hlawka (1961)**, non am liorable)

avec $V(f)$ = variation au sens de Hardy et Krause, et

$$\left| \int_{\mathbb{I}_d} f(\mathbf{u}) d\mathbf{u} - \frac{1}{n} \sum_{i=1}^n f(\mathbf{x}_i) \right| \leq \begin{cases} V(f) C_d \frac{(\log n)^{d-1}}{n} \\ \text{pour } \mathbf{X}_n = \text{ensemble de Hammersley} \\ V(f) C'_d \frac{(\log n)^d}{n} \\ \text{pour } \mathbf{X}_n = n \text{ premiers points} \\ \text{d'une suite de Halton (par ex.) } \mathbf{X}_\infty \end{cases}$$

voir § 2.4

➤ Erreur $\searrow$ plus vite que pour MC, mais quelle valeur de C_d ?

➤ Ensemble $\mathbf{X}_n$ de n points $\neq n$ premiers  l ments

d'une suite infinie de points $\mathbf{X}_\infty$

2.3 Critères de discrédance

Difficulté : $D_n^*(\mathbf{X}_n)$ et $D_n(\mathbf{X}_n)$ sont difficiles à calculer

⇒ (nombreuses!) autres définitions de discrédances

2.3 Critères de discrédance

Difficulté : $D_n^*(\mathbf{X}_n)$ et $D_n(\mathbf{X}_n)$ sont difficiles à calculer

⇒ (nombreuses!) autres définitions de discrédances

On peut souhaiter :

- ① une invariance par permutations des axes, par réflexion par rapport au centre de $\mathbb{I}_d$
- ② une prise en compte de l'uniformité sur les sous-espaces de dim. $< d$
- ③ une interprétation géométrique
- ④ une inégalité de type Koksma-Hlawka
- ⑤ ... et une évaluation facile !

Pour   : changer la norme L_∞ en norme L_p

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

   Expression analytique pour $p = 2$, ok pour   ,    et   , mais pas    et   

Pour   : changer la norme L_∞ en norme L_p

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

   Expression analytique pour $p = 2$, ok pour   ,    et   , mais pas    et   

Pour   : prendre en compte les projections sur toutes les faces de dimension $d' < d$

$$D_n^*(\mathbf{X}_n) \rightsquigarrow \left(\sum_{d'=1}^d \sum_{i_1 < \dots < i_{d'}} \int_{[0,1]^{d'}} |F_n(\{\mathbf{u}\}_{i_1, \dots, i_{d'}}) - F_U(\{\mathbf{u}\}_{i_1, \dots, i_{d'}})|^p d\{\mathbf{u}\}_{i_1, \dots, i_{d'}} \right)^{1/p}$$

   Expression analytique pour $p = 2$, ok pour   ,   ,    et   , mais pas   

Pour  : changer la norme L_∞ en norme L_p

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

  Expression analytique pour $p = 2$, ok pour  ,   et  , mais pas   et  

Pour  : prendre en compte les projections sur toutes les faces de dimension $d' < d$

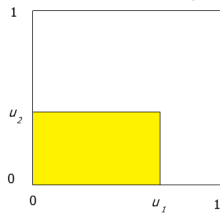
$$D_n^*(\mathbf{X}_n) \rightsquigarrow$$

$$\left(\sum_{d'=1}^d \sum_{i_1 < \dots < i_{d'}} \int_{[0,1]^{d'}} |F_n(\{\mathbf{u}\}_{i_1, \dots, i_{d'}}) - F_U(\{\mathbf{u}\}_{i_1, \dots, i_{d'}})|^p d\{\mathbf{u}\}_{i_1, \dots, i_{d'}} \right)^{1/p}$$

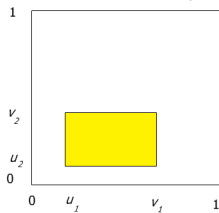
  Expression analytique pour $p = 2$, ok pour  ,  ,   et  , mais pas  

Pour  : changer la forme des ensembles $\mathbb{B}$ dans le calcul de la discr pance

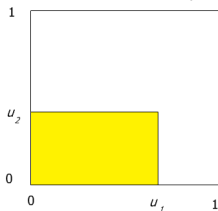
$$D_n^*(\mathbf{X}_n): \mathbb{B} = \prod_{\ell=1}^d [0, u_\ell)$$



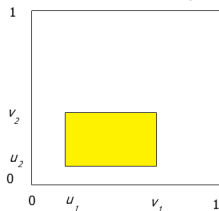
$$D_n(\mathbf{X}_n): \mathbb{B} = \prod_{\ell=1}^d [u_\ell, v_\ell)$$



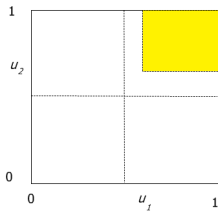
$$D_n^*(\mathbf{X}_n): \mathbb{B} = \prod_{\ell=1}^d [0, u_\ell)$$



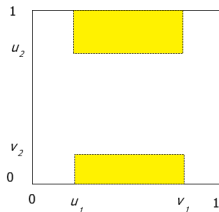
$$D_n(\mathbf{X}_n): \mathbb{B} = \prod_{\ell=1}^d [u_\ell, v_\ell)$$



Discrédance centrée: considérer le sommet du cube (dimension d') le plus proche de $\mathbf{u}$



Discrédance enveloppée (*wrap-around*): considérer $[u_i, v_i]$ si $u_i \leq v_i$ et $[u_i, 1] \cup [0, v_i]$ sinon



Ça se calcule...

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

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

(différentiables en $\mathbf{X}_n$) voir Hickernell (1998a,b); Fang and Ma (2001)
et ça peut donc se minimiser (Fang and Ma, 2001; Fang et al., 2003, 2005)

Ça se calcule...

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

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

(différentiables en $\mathbf{X}_n$) voir Hickernell (1998a,b); Fang and Ma (2001)
 et ça peut donc se minimiser (Fang and Ma, 2001; Fang et al., 2003, 2005)

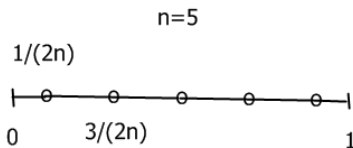
⇒ Il est plus simple de générer des suites de points à faible discrédance...

2.4 Suites à faible discrédance (Niederreiter, 1992, Chap. 3)

$\mathbf{X}_n$ à faible discrédance (*low discrepancy*) si $D_n^*(\mathbf{X}_n^*)$ petit
(ou $D_n(\mathbf{X}_n^*)$, ou autre discrédance)

En dimension 1 :

$\mathbf{X}_n^*$ tel que $x_i = \frac{2i-1}{2n}$,
 $i = 1, \dots, n$



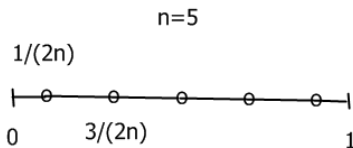
donne $n \times D_n^*(\mathbf{X}_n^*) = 1/2$ pour tout n

2.4 Suites à faible discrédance (Niederreiter, 1992, Chap. 3)

$\mathbf{X}_n$ à faible discrédance (*low discrepancy*) si $D_n^*(\mathbf{X}_n^*)$ petit
(ou $D_n(\mathbf{X}_n^*)$, ou autre discrédance)

En dimension 1 :

$\mathbf{X}_n^*$ tel que $x_i = \frac{2i-1}{2n}$,
 $i = 1, \dots, n$



donne $n \times D_n^*(\mathbf{X}_n^*) = 1/2$ pour tout n

mais on ne peut pas construire une suite $\mathbf{X}_\infty = (x_1, x_2, x_3 \dots)$ telle que pour n élément consécutifs on ait $n \times D_n^*(\mathbf{X}_\infty) = \mathcal{O}(1)$

Illustration : partons de $n = 4$, $\mathbf{X}_4^* = (1/8, 3/8, 5/8, 7/8)$

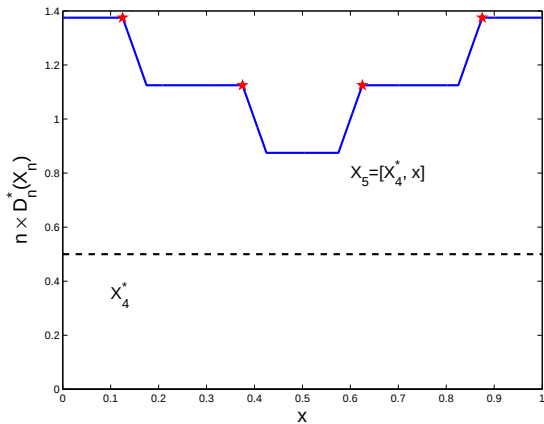


Illustration : partons de $n = 4$, $\mathbf{X}_4^* = (1/8, 3/8, 5/8, 7/8)$

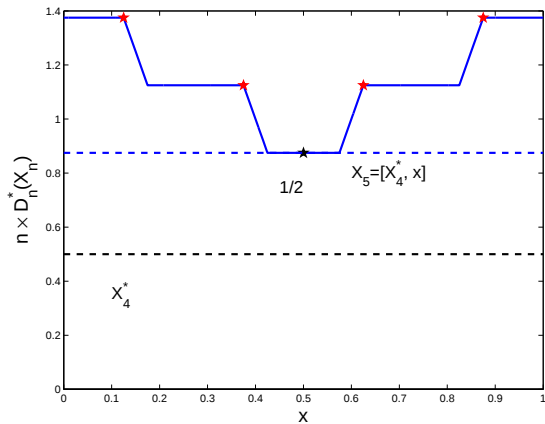


Illustration : partons de $n = 4$, $\mathbf{X}_4^* = (1/8, 3/8, 5/8, 7/8)$

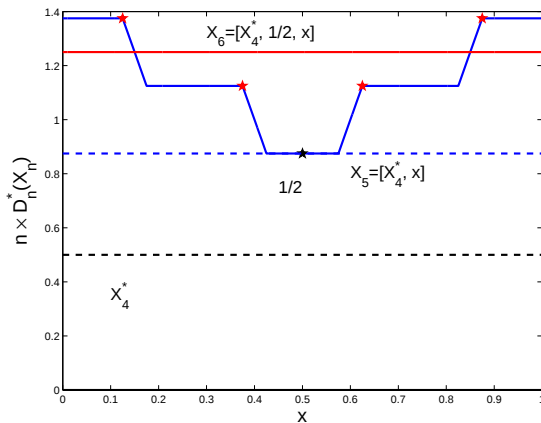
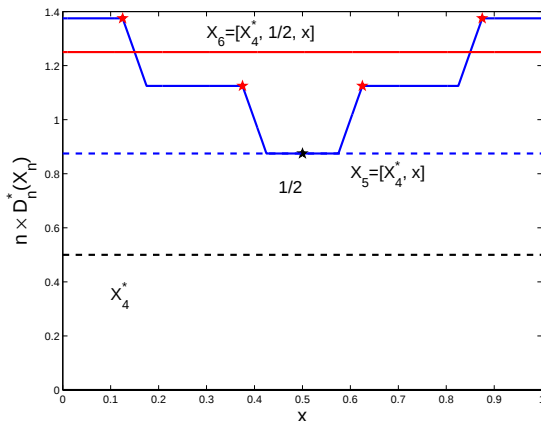


Illustration : partons de $n = 4$, $\mathbf{X}_4^* = (1/8, 3/8, 5/8, 7/8)$



$n \times D_n^*(\mathbf{X}_\infty)$ fluctue forc ment : $D_n^*(\mathbf{X}_\infty) > c \frac{\log n}{n}$ infiniment souvent
(meilleure constante connue $c = 0.06$)

On sait construire des suites $\mathbf{X}_\infty$ telles que $D_n^*(\mathbf{X}_\infty) = \mathcal{O}\left(\frac{\log n}{n}\right)$

- Parties fractionnaires (surtout pour $d = 1$)
  n fix  $\leftrightarrow$ treillis, r seaux de points (*Lattices*)
- Suites de van der Corput ($d = 1$)
- $\Rightarrow$ Suites de Halton ($d > 1$)
- (t, m, d) -nets et suites- (t, d) (Sobol Faure)

- Parties fractionnaires (surtout pour $d = 1$)
  n fix  $\leftrightarrow$ treillis, r seaux de points (*Lattices*)
- Suites de van der Corput ($d = 1$)
- $\Rightarrow$ Suites de Halton ($d > 1$)
- (t, m, d) -nets et suites- (t, d) (Sobol Faure)

$\Rightarrow$ van der Corput

Parties fractionnaires

Pour $d = 1$: $x_k \triangleq \{kz\} = kz - \lfloor kz \rfloor$, $k = 1, 2, \dots$, avec z irrationnel donn 
(partie fractionnaire de kz)

Par exemple $z = \varphi = (\sqrt{5} + 1)/2 \simeq 1,618034 = \text{nombre d'or}$

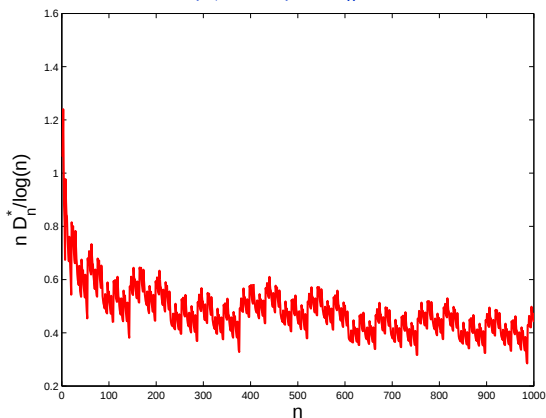
Parties fractionnaires

Pour $d = 1$: $x_k \triangleq \{kz\} = kz - \lfloor kz \rfloor$, $k = 1, 2, \dots$, avec z irrationnel donn 

(partie fractionnaire de kz)

Par exemple $z = \varphi = (\sqrt{5} + 1)/2 \simeq 1,618034 =$ nombre d'or

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

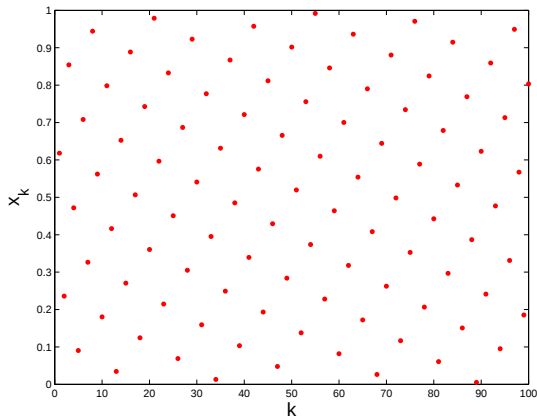


Parties fractionnaires

Pour $d = 1$: $x_k \triangleq \{kz\} = kz - \lfloor kz \rfloor$, $k = 1, 2, \dots$, avec z irrationnel donné
(partie fractionnaire de kz)

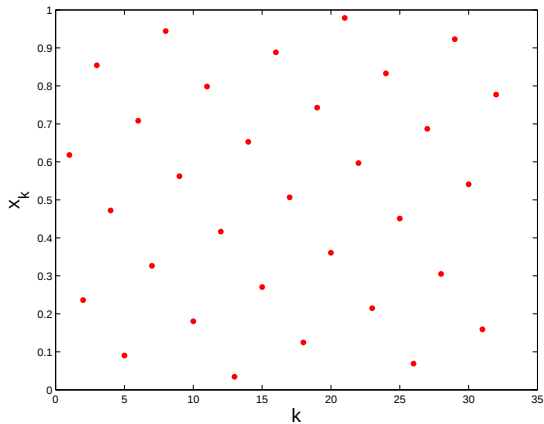
Par exemple $z = \varphi = (\sqrt{5} + 1)/2 \simeq 1,618034 =$ nombre d'or

100 premiers points x_k



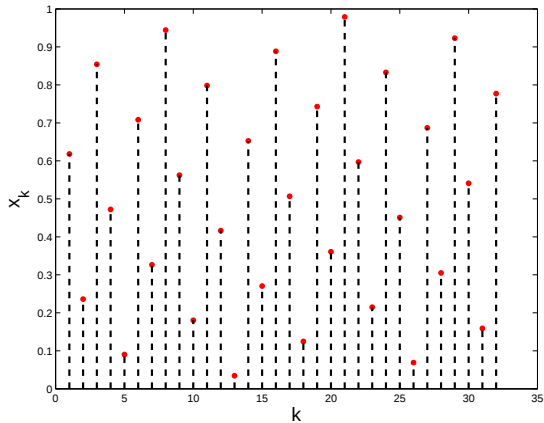
Pour $d > 1$:

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



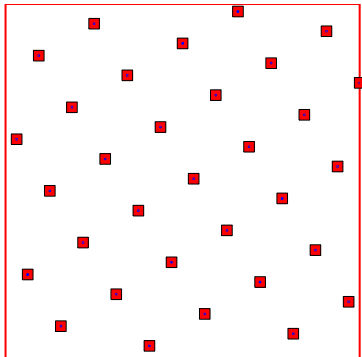
Pour $d > 1$:

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

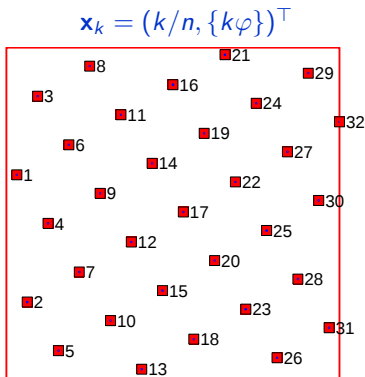


Pour $d > 1$:

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



Pour $d > 1$:



⇒ Remplacer k/n (croissance monotone et valable seulement pour $k = 1, \dots, n$) par $\{kz\}$, z irrationnel $\rightarrow \mathbf{x}_k \in [0, 1]^2$, $k = 1, 2, 3, \dots$

⇒ on répète... $\rightarrow$ récursivement $\mathbf{x}_k \in [0, 1]^d$, d quelconque

Prendre $\mathbf{x}_k \triangleq \{k\mathbf{z}\}$, $\mathbf{z}$ vecteur irrationnel de $\mathbb{R}^d$

(avec composantes ind pendantes sur les rationnels: $\mathbf{q}^\top \mathbf{z} \neq 0, \forall \mathbf{q} \in \mathbb{Q}^d$)

⇒ suite $\mathbf{X}_\infty$ distribu e uniform ment dans $[0, 1]^d$

(Kuipers and Niederreiter, 1974, p. 48)

⇒ $\forall \epsilon > 0, D_n(\mathbf{X}_\infty) = \mathcal{O}\left(\frac{(1+\log n)^{d+1+\epsilon}}{n}\right)$ pour presque tout $\mathbf{z}$

Int ressant, mais ne dit pas quel $\mathbf{z}$ prendre

Prendre $\mathbf{x}_k \triangleq \{k\mathbf{z}\}$, $\mathbf{z}$ vecteur irrationnel de $\mathbb{R}^d$

(avec composantes indépendantes sur les rationnels: $\mathbf{q}^\top \mathbf{z} \neq 0, \forall \mathbf{q} \in \mathbb{Q}^d$)

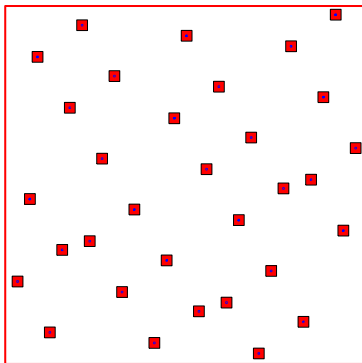
⇒ suite $\mathbf{X}_\infty$ distribuée uniformément dans $[0, 1]^d$

(Kuipers and Niederreiter, 1974, p. 48)

⇒ $\forall \epsilon > 0, D_n(\mathbf{X}_\infty) = \mathcal{O}\left(\frac{(1+\log n)^{d+1+\epsilon}}{n}\right)$ pour presque tout $\mathbf{z}$

Intéressant, mais ne dit pas quel $\mathbf{z}$ prendre

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



Prendre $\mathbf{x}_k \triangleq \{k\mathbf{z}\}$, $\mathbf{z}$ vecteur irrationnel de $\mathbb{R}^d$

(avec composantes indépendantes sur les rationnels: $\mathbf{q}^\top \mathbf{z} \neq 0, \forall \mathbf{q} \in \mathbb{Q}^d$)

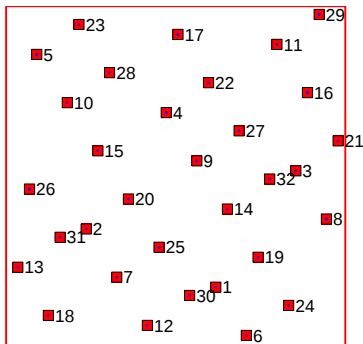
⇒ suite $\mathbf{X}_\infty$ distribuée uniformément dans $[0, 1]^d$

(Kuipers and Niederreiter, 1974, p. 48)

⇒ $\forall \epsilon > 0, D_n(\mathbf{X}_\infty) = \mathcal{O}\left(\frac{(1+\log n)^{d+1+\epsilon}}{n}\right)$ pour presque tout $\mathbf{z}$

Intéressant, mais ne dit pas quel $\mathbf{z}$ prendre

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



Treillis, r seaux de points (*Lattices*) : $\mathbf{z} = \mathbf{g}/n$, avec $\mathbf{g} \in \mathbb{Z}^d$

- ⇒ $\mathbf{x}_k \triangleq \left\{ \frac{k}{n} \mathbf{g} \right\}$ (avec r p titions si $k \geq n$!)
 n points $\neq$ si $\text{pgcd}(g_1, \dots, g_d, n) = 1$
 n points $\neq$ pour chaque coordonn e si $\text{pgcd}(g_i, n) = 1$ pour tout i
- ⇒ Espacements r guliers entre les points ($\approx$ grilles)
permet de prendre en compte la r gularit  de $f(\cdot)$ dans le calcul de borne sur l'erreur d'int gration

Treillis, r seaux de points (*Lattices*) : $\mathbf{z} = \mathbf{g}/n$, avec $\mathbf{g} \in \mathbb{Z}^d$

- $\Rightarrow \mathbf{x}_k \triangleq \left\{ \frac{k}{n} \mathbf{g} \right\}$ (avec r p titions si $k \geq n$!)
 - n points $\neq$ si $\text{pgcd}(g_1, \dots, g_d, n) = 1$
 - n points $\neq$ pour chaque coordonn e si $\text{pgcd}(g_i, n) = 1$ pour tout i
- $\Rightarrow$ Espacements r guliers entre les points ($\approx$ grilles)
 - permet de prendre en compte la r gularit  de $f(\cdot)$ dans le calcul de borne sur l'erreur d'int gration

Pour $d = 2$, un tr s bon choix est $n = F_m$ pour $\mathbf{g} = (1, F_{m-1})^\top$,
 avec (F_m) = suite de Fibonacci : $F_1 = F_2 = 1$, $F_{k+1} = F_k + F_{k-1}$, $k \geq 2$

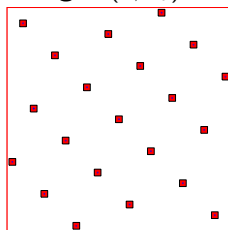
Comme $F_{m-1}/F_m \rightarrow 1/\varphi$ pour $m \rightarrow \infty$, construction assez proche de

$$\mathbf{x}_k = \left(\frac{k}{n}, \left\{ \frac{k}{\varphi} \right\} \right)^\top = \left(\frac{k}{n}, \{k\varphi\} \right)^\top, \quad k = 1, \dots, n$$

- Forts liens avec plans optimaux pour r gression de Fourier (sin, cos) (Bates et al., 1996; Riccomagno et al., 1997)
- Il existe des construction (non explicites) ayant de bonnes propri t s (*good lattice points*) ⇨ tables (Maisonneuve, 1972)
- Korobov (1960) sugg re $\mathbf{g} = (1, g, g^2, \dots, g^{d-1})^\top$

- Forts liens avec plans optimaux pour r gression de Fourier (sin, cos) (Bates et al., 1996; Riccomagno et al., 1997)
- Il existe des construction (non explicites) ayant de bonnes propri t s (*good lattice points*) ➡ tables (Maisonneuve, 1972)
- Korobov (1960) sugg re $\mathbf{g} = (1, g, g^2, \dots, g^{d-1})^\top$
- Optimisation du choix de $\mathbf{g}$ dans (par ex.) (Sloan and Walsh, 1990; Sloan and Reztsov, 2002; Nuyens, 2007)

Ex. : $d = 2$, $n = 21 = F_8$
 $\mathbf{g} = (1, F_7)$



$$\Phi_{Mm} = 0.2020$$

$$\Phi_{mM} = 0.2357$$

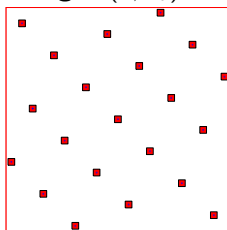
$$D_{Cent, L_2} = 0.0280$$

$$D_{WA, L_2} = 0.0388$$

- Forts liens avec plans optimaux pour r gression de Fourier (sin, cos) (Bates et al., 1996; Riccomagno et al., 1997)
- Il existe des construction (non explicites) ayant de bonnes propri t s (*good lattice points*) ➡ tables (Maisonneuve, 1972)
- Korobov (1960) sugg re $\mathbf{g} = (1, g, g^2, \dots, g^{d-1})^\top$
- Optimisation du choix de $\mathbf{g}$ dans (par ex.) (Sloan and Walsh, 1990; Sloan and Reztsov, 2002; Nuyens, 2007)

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

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



$$\Phi_{Mm} = 0.2020$$

$$\Phi_{mM} = 0.2357$$

$$D_{Cent, L_2} = 0.0280$$

$$D_{WA, L_2} = 0.0388$$

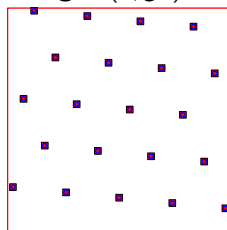
$$\Phi_{Mm} = 0.2302$$

$$\Phi_{mM} = 0.2217$$

$$D_{Cent, L_2} = 0.0536$$

$$D_{WA, L_2} = 0.0633$$

$$\mathbf{g} = (1, g^*)$$



- On ne peut générer que n points :
 - ➡ suite infinie de points dans $[0, 1]^d$ si $\mathbf{x}_k \triangleq \{u_k \mathbf{g}\}$
avec (u_k) une suite (scalaire) à faible discrédance (Hickernell, 1998b)

- On ne peut générer que n points :
- ▢ suite infinie de points dans $[0, 1]^d$ si $\mathbf{x}_k \triangleq \{u_k \mathbf{g}\}$
avec (u_k) une suite (scalaire) à faible discrédance (Hickernell, 1998b)

On a vu seulement la règle de rang 1 (*rank-on rule*)

règle de rang r : $\mathbf{x}_{k_1, \dots, k_r} \triangleq \left\{ \frac{k_1}{n_1} \mathbf{g}_1 + \frac{k_2}{n_2} \mathbf{g}_2 + \dots + \frac{k_r}{n_r} \mathbf{g}_r \right\}$, $k_j \in \{1, \dots, n_j\}$

Règle de copie (*copy rule*) : diviser $[0, 1]^d$ en k^d cubes de longueur d'arête $1/k$,
construire un treillis dans chacun

Suites de van der Corput (1935)

Soit $\mathcal{Z}_b = \{0, 1, \dots, b-1\}$ l'alphabet pour la base $b \geq 2$

(ex., $\mathcal{Z}_2 = \{0, 1\}$, $\mathcal{Z}_3 = \{0, 1, 2\}$)

► Tout $k = 0, 1, \dots, b^m - 1$ peut s' crire comme $k = \sum_{\ell=0}^{m-1} a_\ell b^\ell$
avec m caract res $a_0, a_1, \dots, a_{m-1}$ (d pendant de k)
(c'est- -dire $k = \underline{a_{m-1}a_{m-2} \cdots a_2a_1a_0}_b$)

► On lui associe $\Phi_b(k) = \sum_{\ell=0}^{m-1} a_\ell b^{-(\ell+1)}$

⇒ La suite de van der Corput en base b est d finie par $x_k \triangleq \Phi_b(k)$

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

k	k en base 2	$\phi_b(k)$ en base 2	$\phi_2(k)$
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	$\underline{a_{m-1} \cdots a_1 a_0}_2$	$0, a_0 a_1 \cdots a_{m-1}$	$\sum_{\ell=0}^{m-1} a_\ell 2^{-(\ell+1)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

$$\Rightarrow nD_n^*(\mathbf{X}_\infty) = nD_n(\mathbf{X}_\infty) \leq 1 + \frac{\log n}{\log 2}$$

base $b = 3$

k	k en 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	$8/9$
$\vdots$	$\vdots$	$\vdots$

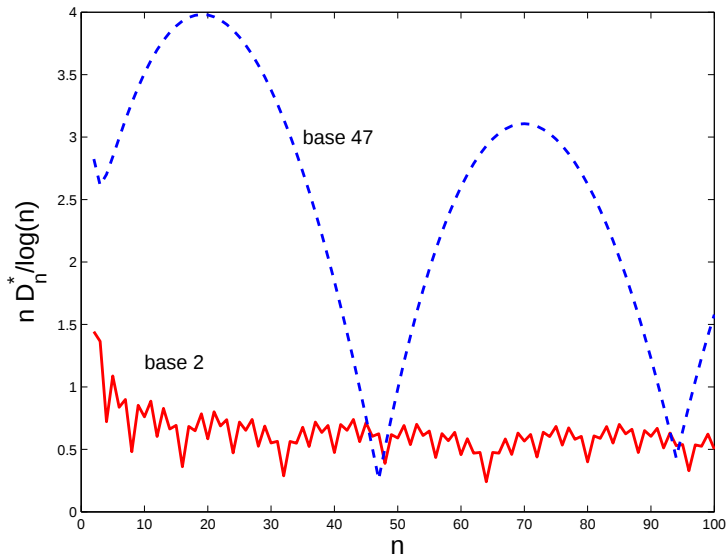
base $b = 3$

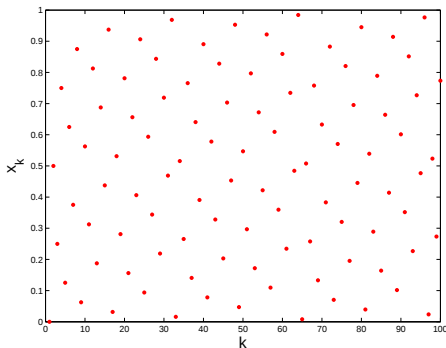
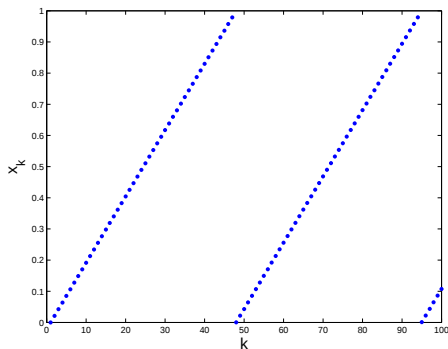
k	k en base 3	$\phi_3(k)$
0	0	0
1	1	1/3
2	2	2/3
3	10	1/9
4	11	4/9
5	12	7/9
6	20	2/9
7	21	5/9
8	22	8/9
$\vdots$	$\vdots$	$\vdots$

Choix particulier de b + permutation adéquate de $\mathcal{L}_b$

⇒ meilleures performances connues pour $\limsup_{n \rightarrow \infty} nD_n^*/\log(n)$ ($b = 12$) et $\limsup_{n \rightarrow \infty} nD_n/\log(n)$ ($b = 36$) (travaux de H. Faure (1977–20xx))

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



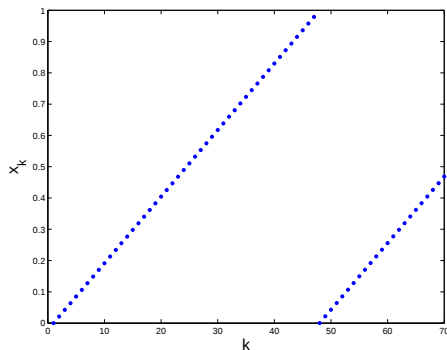
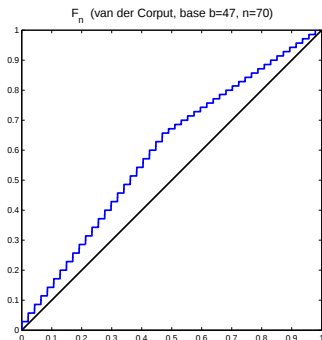
van der Corput : $x_1, x_2, \dots, x_{100}$ $b = 2$  $b = 47, x_1, x_2, \dots$ 

...et de plus, le “remplissage” de $[0, 1]$ se fait dans un ordre particulier

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

$b = 47$, fonct. répartition empirique

$b = 47, x_1, x_2 \dots$



... et de plus, le “remplissage” de $[0,1]$ se fait dans un ordre particulier

Pour $d > 1$: on travaille s par ment sur chaque composante

$b_1, b_2 \dots$ des entiers tels que $\text{pgcd}(b_i, b_j) = 1$ pour tout $i \neq j$

(en pratique, les premiers nombres premiers:

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

Pour $d > 1$: on travaille s par ment sur chaque composante

$b_1, b_2 \dots$ des entiers tels que $\text{pgcd}(b_i, b_j) = 1$ pour tout $i \neq j$

(en pratique, les premiers nombres premiers:

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

► Hammersley : $\mathbf{x}_k \triangleq (k/n, \phi_{b_1}(k), \dots, \phi_{b_{d-1}}(k))^T$, $k = 1, \dots, n$
(n est fix )

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

Pour $d > 1$: on travaille s par ment sur chaque composante

$b_1, b_2 \dots$ des entiers tels que $\text{pgcd}(b_i, b_j) = 1$ pour tout $i \neq j$

(en pratique, les premiers nombres premiers:

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

► **Hammersley** : $\mathbf{x}_k \triangleq (k/n, \phi_{b_1}(k), \dots, \phi_{b_{d-1}}(k))^T, k = 1, \dots, n$
(n est fix )

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

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

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

Pour $d > 1$: on travaille séparément sur chaque composante

$b_1, b_2 \dots$ des entiers tels que $\text{pgcd}(b_i, b_j) = 1$ pour tout $i \neq j$

(en pratique, les premiers nombres premiers:

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

➤ **Hammersley** : $\mathbf{x}_k \triangleq (k/n, \phi_{b_1}(k), \dots, \phi_{b_{d-1}}(k))^T, k = 1, \dots, n$
(n est fixé)

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

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

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

On pense que dans tous les cas (prouvé seulement pour $d = 1, 2$)

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

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

Les vitesses de décroissance de D_n^* sont donc optimales pour **Hammersley** et **Halton**

Pour $d > 1$: on travaille séparément sur chaque composante

$b_1, b_2 \dots$ des entiers tels que $\text{pgcd}(b_i, b_j) = 1$ pour tout $i \neq j$
(en pratique, les premiers nombres premiers:
 $2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47 \dots$)

➤ **Hammersley** : $\mathbf{x}_k \triangleq (k/n, \phi_{b_1}(k), \dots, \phi_{b_{d-1}}(k))^T, k = 1, \dots, n$
(n est fixé)

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

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

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

On pense que dans tous les cas (prouvé seulement pour $d = 1, 2$)

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

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

Les vitesses de décroissance de D_n^* sont donc optimales pour **Hammersley** et **Halton** ... mais quid de la constante A_d ?

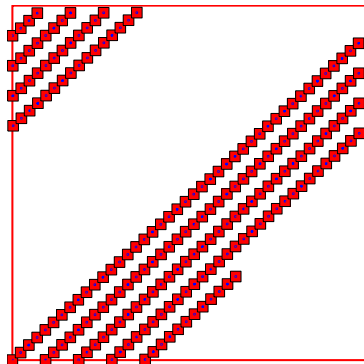
On montre que $\lim_{d \rightarrow \infty} \frac{\log A_d}{d \log d} = 1$

($A_d \nearrow$ superexponentiellement vite avec d !)

On montre que $\lim_{d \rightarrow \infty} \frac{\log A_d}{d \log d} = 1$

($A_d \nearrow$ superexponentiellement vite avec d !)

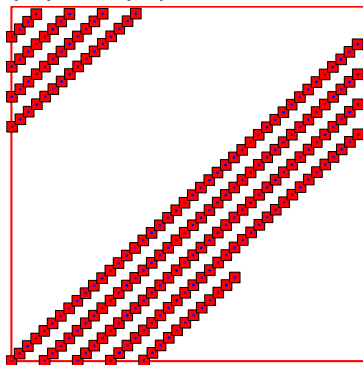
⇒ **Discrédance pas très bonne quand d est grand**
 suite de Halton, $d = 15$ ($\Rightarrow b_{15} = 47$) :
 $\{\mathbf{x}_k\}_{14}$ et $\{\mathbf{x}_k\}_{15}$, $k = 1, \dots, 200$



On montre que $\lim_{d \rightarrow \infty} \frac{\log A_d}{d \log d} = 1$

($A_d \nearrow$ superexponentiellement vite avec d !)

⇒ **Discrédance pas très bonne quand d est grand**
 suite de Halton, $d = 15$ ($\Rightarrow b_{15} = 47$) :
 $\{\mathbf{x}_k\}_{14}$ et $\{\mathbf{x}_k\}_{15}$, $k = 1, \dots, 200$



... pas trop mal dans le plan $\{\mathbf{x}_k\}_{d_1}, \{\mathbf{x}_k\}_{d_2}$ si $n = b_{d_1} b_{d_2}$
 (avec $b_d \approx d(\log d + \log \log d)$)

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

► suites- (t, d)

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

► suites-(t, d)

Motivation : combattre le probl me de $A_d \nearrow \infty$ quand $d \rightarrow \infty$
pour les suites de Halton

Pour une base b , on consid re un pav   l mentaire

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

o  les q_j et a_j sont des entiers, $0 \leq q_j$ et $0 \leq a_j \leq b^{q_j} - 1$

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

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

► suites-(t, d)

Motivation : combattre le probl me de $A_d \nearrow \infty$ quand $d \rightarrow \infty$
pour les suites de Halton

Pour une base b , on consid re un pav   l mentaire

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

o  les q_j et a_j sont des entiers, $0 \leq q_j$ et $0 \leq a_j \leq b^{q_j} - 1$

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

Objectif : mettre des points dans tous les pav s  l mentaires
(en consid rant tous les d coupages possibles en pav s  l mentaires !)

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

► suites- (t, d)

Motivation : combattre le probl me de $A_d \nearrow \infty$ quand $d \rightarrow \infty$
pour les suites de Halton

Pour une base b , on consid re un pav   l mentaire

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

o  les q_j et a_j sont des entiers, $0 \leq q_j$ et $0 \leq a_j \leq b^{q_j} - 1$

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

Objectif : mettre des points dans tous les pav s  l mentaires
(en consid rant tous les d coupages possibles en pav s  l mentaires !)

Plus pr cis ment : pour $0 \leq t \leq m$, un (t, m, d) -net en base b poss de $n = b^m$ points, tels que chaque pav   l mentaire de volume b^{t-m} contienne b^t points

Exemple : $(0, 2, 2)$ -net en base 2 ($b = 2$, $d = 2$, $m = 2$, $t = 0$)

⇒ $n = b^m = 4$, $b^0 = 1$ point dans chaque pav  de volume $b^{t-m} = 1/4$

Exemple : $(0, 2, 2)$ -net en base 2 ($b = 2, d = 2, m = 2, t = 0$)

⇒ $n = b^m = 4$, $b^0 = 1$ point dans chaque pav  de 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	$a_1 \in \{0, 1\}, a_2 \in \{0, 1\}$

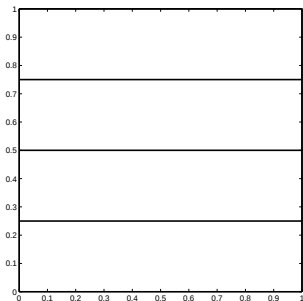
Exemple : $(0, 2, 2)$ -net en base 2 ($b = 2$, $d = 2$, $m = 2$, $t = 0$)

⇒ $n = b^m = 4$, $b^0 = 1$ point dans chaque pavé de volume $b^{t-m} = 1/4$

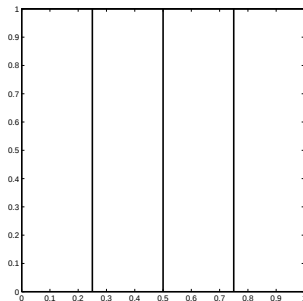
$$\sum_{j=1}^d q_j = m - t = 2 \Rightarrow q_j \in \{0, 1, 2\}$$

	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	$a_1 \in \{0, 1\}, a_2 \in \{0, 1\}$

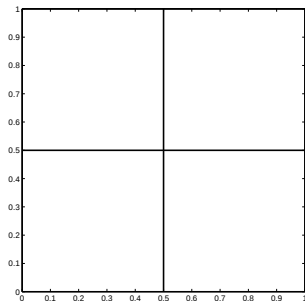
cas (i)



cas (ii)



cas (iii)



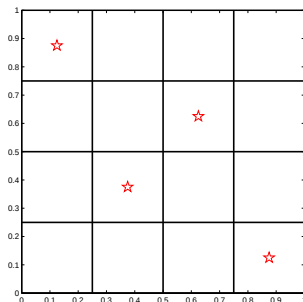
Exemple : (0, 2, 2)-net en base 2 ($b = 2$, $d = 2$, $m = 2$, $t = 0$)

⇒ $n = b^m = 4$, $b^0 = 1$ point dans chaque pav  de 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	$a_1 \in \{0, 1\}, a_2 \in \{0, 1\}$

un (0, 2, 2)-net en base 2



Quelques propri t s :

- $t = m$: cas trivial $\Rightarrow b^t = b^m = n$ points (tous) dans le pav  de volume $b^{t-m} = 1$

Quelques propriétés :

- $t = m$: cas trivial $\Rightarrow b^t = b^m = n$ points (tous) dans le pavé de volume $b^{t-m} = 1$
- (t, m, d) -net en base $b \Rightarrow (t', m, d)$ -net en base b pour $t \leq t' \leq m$
 $\Rightarrow$ la difficulté augmente quand t diminue !

Quelques propri t s :

- $t = m$: cas trivial $\Rightarrow b^t = b^m = n$ points (tous) dans le pav  de volume $b^{t-m} = 1$
- (t, m, d) -net en base $b \Rightarrow (t', m, d)$ -net en base b pour $t \leq t' \leq m$
 $\Rightarrow$ la difficult  augmente quand t diminue !
- $t = 0, m = 1$, base $b \Rightarrow n = b^m = b$ points, $b^t = 1$ point par pav   l mentaire
 $\sum_{j=1}^d q_j = 1 \Rightarrow$ un seul $q_j \neq 0$, $q_j = 1 \Rightarrow a_j \in \{0, 1, \dots, b-1\}$
 $\Rightarrow (0, 1, d)$ -net en base $b =$ hypercube latin   b points

Quelques propri t s :

- $t = m$: cas trivial $\Rightarrow b^t = b^m = n$ points (tous) dans le pav  de volume $b^{t-m} = 1$
- (t, m, d) -net en base $b \Rightarrow (t', m, d)$ -net en base b pour $t \leq t' \leq m$
 $\Rightarrow$ la difficult  augmente quand t diminue !
- $t = 0, m = 1$, base $b \Rightarrow n = b^m = b$ points, $b^t = 1$ point par pav   l mentaire
 $\sum_{j=1}^d q_j = 1 \Rightarrow$ un seul $q_j \neq 0$, $q_j = 1 \Rightarrow a_j \in \{0, 1, \dots, b-1\}$
 $\Rightarrow (0, 1, d)$ -net en base $b =$ hypercube latin   b points
- $(0, 2, d)$ -net en base $b =$ OA (*Orthogonal Array*) + hypercube latin   $n = b^2$ points: exemple pr c dent avec $b = 2$ et $d = 2$, mais n'existe pas toujours. . .
- pas de $(0, m, d)$ -net en base b pour $d > b + 1$ (Niederreiter, 1992, p. 62)

$\mathbf{X}_\infty = \boxed{\text{suite-}(t, d)}$ en base b si pour tout $k \geq 0$ et tout $m \geq t$
les b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \dots, \mathbf{x}_{(k+1)b^m-1}\}$ forment un (t, m, d) -net

$\mathbf{X}_\infty = \boxed{\text{suite-}(t, d)}$ en base b si pour tout $k \geq 0$ et tout $m \geq t$
les b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \dots, \mathbf{x}_{(k+1)b^m-1}\}$ forment un (t, m, d) -net

Quelques propri t s :

- Surtout int ressant pour t petit

$\mathbf{X}_\infty = \boxed{\text{suite-}(t, d)}$ en base b si pour tout $k \geq 0$ et tout $m \geq t$
les b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \dots, \mathbf{x}_{(k+1)b^m-1}\}$ forment un (t, m, d) -net

Quelques propri t s :

- Surtout int ressant pour t petit
- pas de suite- $(0, d)$ en base b pour $d > b$ (Niederreiter, 1992, p. 62)
suite- $(0, d)$ en base $b \geq d$ avec b premier = suite de Faure (1982)

$\mathbf{X}_\infty = \boxed{\text{suite-}(t, d)}$ en base b si pour tout $k \geq 0$ et tout $m \geq t$
les b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \dots, \mathbf{x}_{(k+1)b^m-1}\}$ forment un (t, m, d) -net

Quelques propri t s :

- Surtout int ressant pour t petit
- pas de suite- $(0, d)$ en base b pour $d > b$ (Niederreiter, 1992, p. 62)
suite- $(0, d)$ en base $b \geq d$ avec b premier = suite de Faure (1982)
- Suite de van der Corput en base $b = \text{suite-}(0, 1)$ en base b

$\mathbf{X}_\infty = \boxed{\text{suite-}(t, d)}$ en base b si pour tout $k \geq 0$ et tout $m \geq t$ les b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \dots, \mathbf{x}_{(k+1)b^m-1}\}$ forment un (t, m, d) -net

Quelques propri t s :

- Surtout int ressant pour t petit
- pas de suite- $(0, d)$ en base b pour $d > b$ (Niederreiter, 1992, p. 62)
suite- $(0, d)$ en base $b \geq d$ avec b premier = suite de Faure (1982)
- Suite de van der Corput en base $b = \text{suite-}(0, 1)$ en base b
- Construction un peu compliqu e. . .
Pour $b = 2$, suites de Sobol (1967), d quelconque — avec un t plus petit (pour $d \geq 8$) pour les suites de Niederreiter (1992)

$\mathbf{X}_\infty = \boxed{\text{suite-}(t, d)}$ en base b si pour tout $k \geq 0$ et tout $m \geq t$ les b^m points $\{\mathbf{x}_{kb^m}, \mathbf{x}_{kb^m+1}, \dots, \mathbf{x}_{(k+1)b^m-1}\}$ forment un (t, m, d) -net

Quelques propriétés :

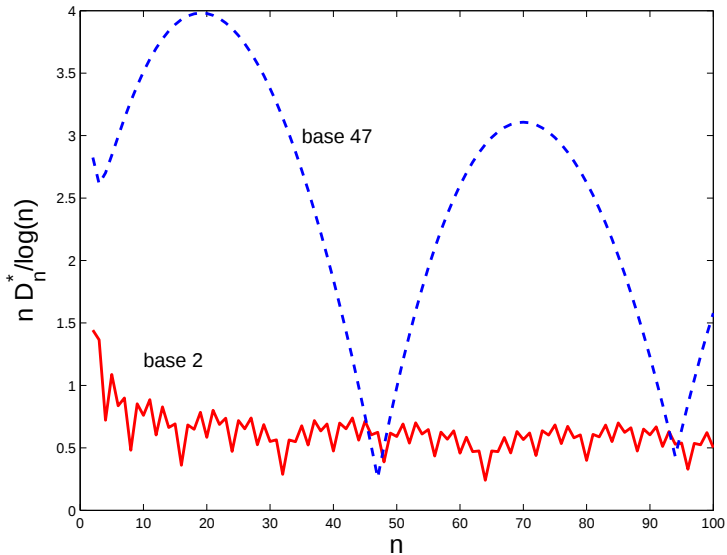
- Surtout intéressant pour t petit
- pas de suite- $(0, d)$ en base b pour $d > b$ (Niederreiter, 1992, p. 62)
suite- $(0, d)$ en base $b \geq d$ avec b premier = suite de Faure (1982)
- Suite de van der Corput en base $b = \text{suite-}(0, 1)$ en base b
- Construction un peu compliquée...
Pour $b = 2$, suites de Sobol (1967), d quelconque — avec un t plus petit (pour $d \geq 8$) pour les suites de Niederreiter (1992)

► suites- (t, d) avec t et b aussi petits que possible et discrédance

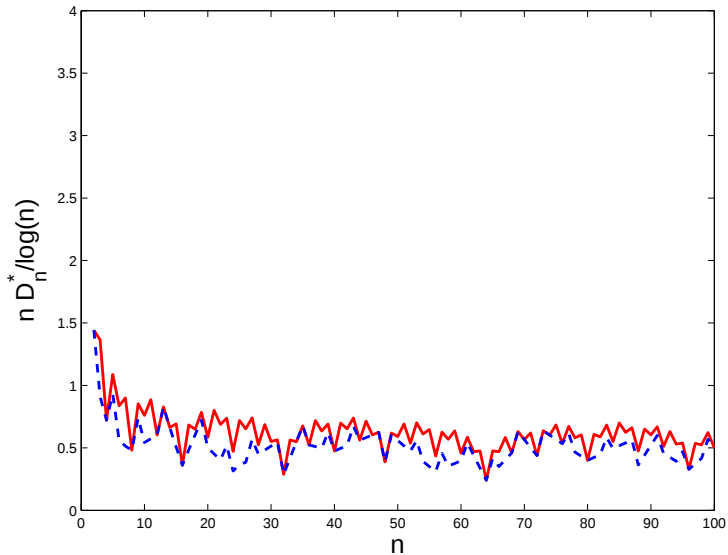
$$D_n^*(\mathbf{X}_\infty) \leq C_d \frac{(\log n)^d}{n} + \mathcal{O}\left(\frac{(\log n)^{d-1}}{n}\right)$$

“ $\approx$ Halton”... mais ici $C_d \searrow 0$ superexponentiellement quand $d \rightarrow \infty$!

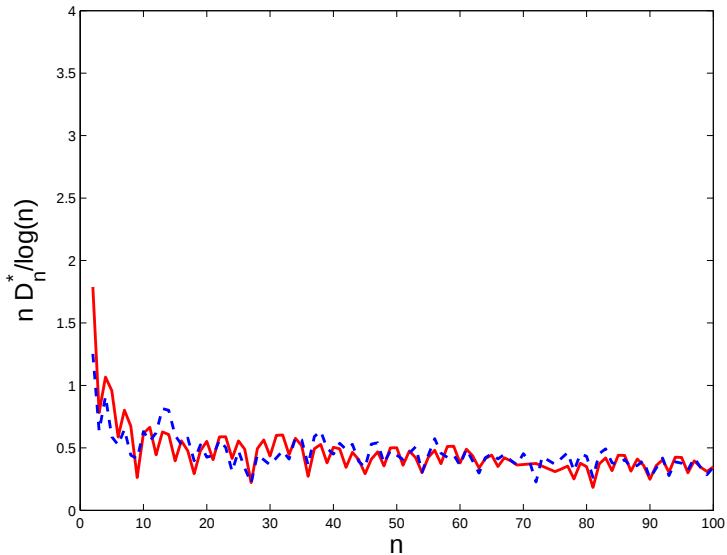
Halton : $(n/\log n) \times D_n^*$ pour $\{\mathbf{X}_n\}_1$ et $\{\mathbf{X}_n\}_{15}$



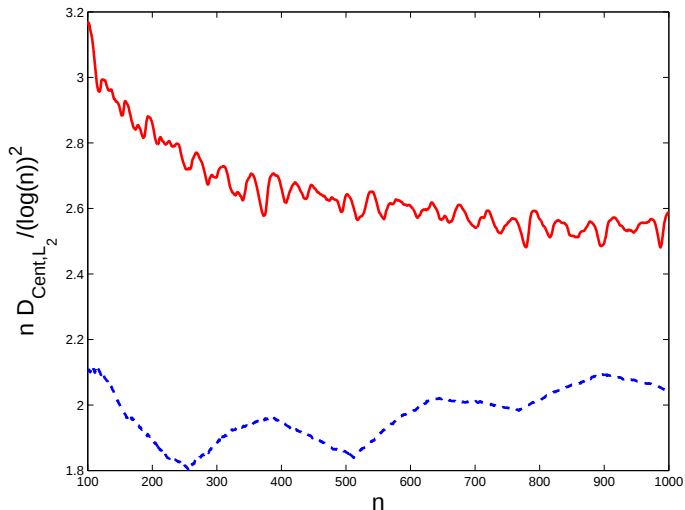
Sobol : $(n/\log n) \times D_n^*$ pour $\{\mathbf{X}_n\}_1$ et $\{\mathbf{X}_n\}_{15}$



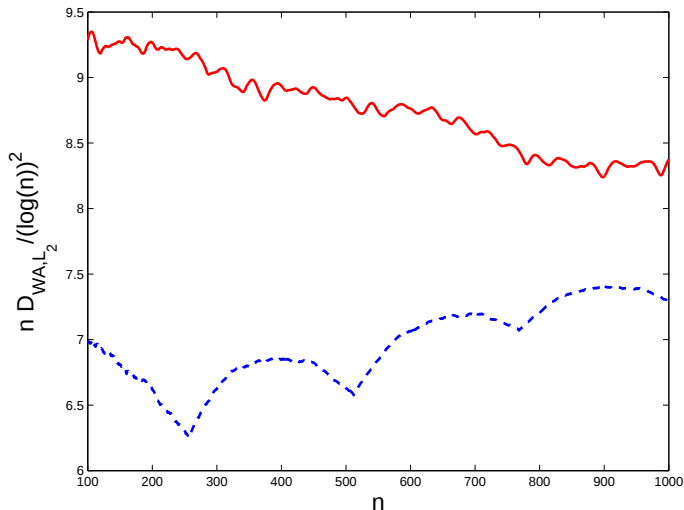
Faure : $(n/\log n) \times D_n^*$ pour $\{\mathbf{X}_n\}_1$ et $\{\mathbf{X}_n\}_{15}$



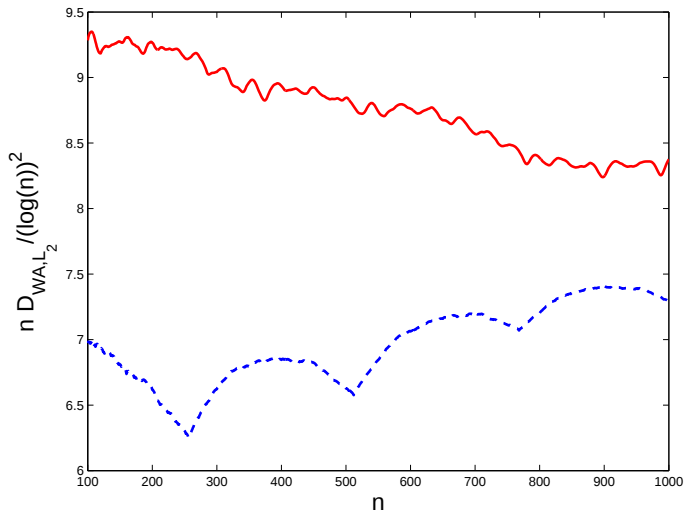
$[n/(\log n)^2] \times D_{Cent, L_2}(\mathbf{X}_n)$ pour **Halton** et **Sobol** ($d = 15$)
 $(/(\log n)^2)$, mais pas de normalisation évidente pour D_{Cent, L_2}



$[n/(\log n)^2] \times D_{WA,L_2}(\mathbf{X}_n)$ pour **Halton** et **Sobol** ($d = 15$)
 $(/(\log n)^2)$, mais pas de normalisation  vidente pour D_{WA,L_2})

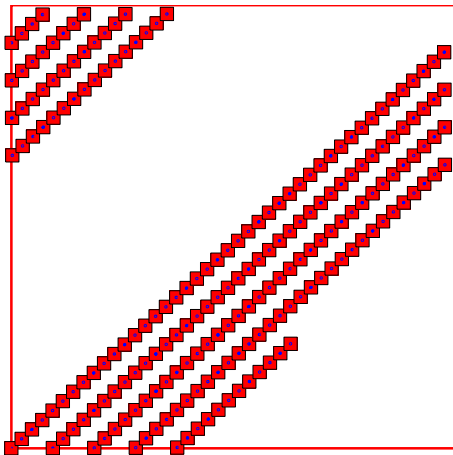


$[n/(\log n)^2] \times D_{WA,L_2}(\mathbf{X}_n)$ pour **Halton** et **Sobol** ($d = 15$)
 $(/(\log n)^2)$, mais pas de normalisation  vidente pour D_{WA,L_2}

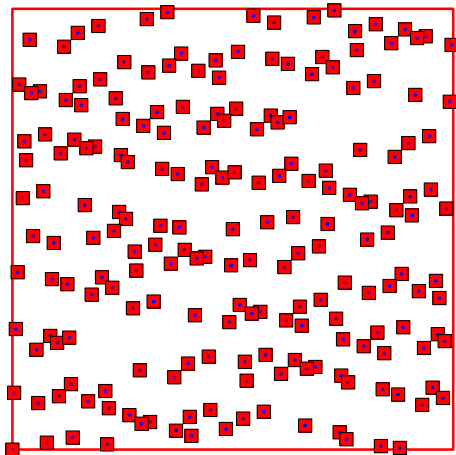


► $n = 1\,000$ est petit : le “cube” $[0, 1]^{15}$ a $2^{15} = 32\,768$ sommets !

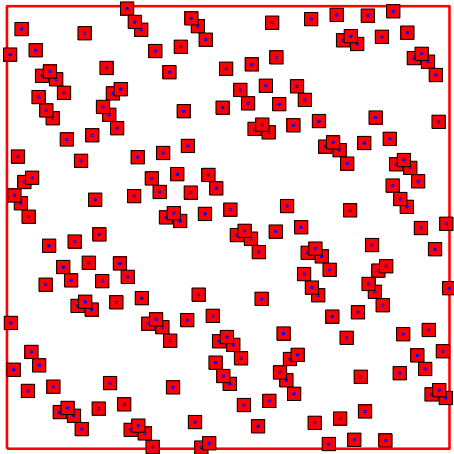
Halton, $d = 15$: $\{\mathbf{x}_k\}_{14}$ et $\{\mathbf{x}_k\}_{15}$,
 $k = 1, \dots, 200$



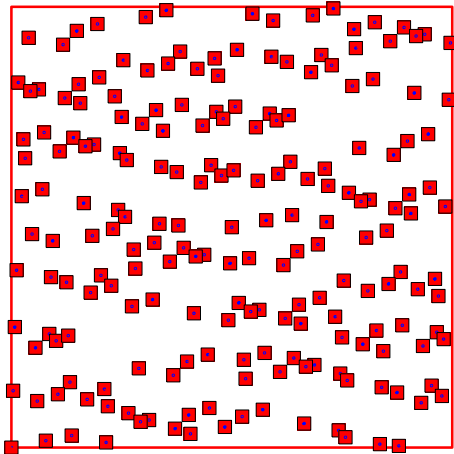
Sobol : $d = 15$, $\{\mathbf{x}_k\}_{14}$ et $\{\mathbf{x}_k\}_{15}$,
 $k = 1, \dots, 200$



Faure, $d = 15$: $\{\mathbf{x}_k\}_{14}$ et $\{\mathbf{x}_k\}_{15}$,
 $k = 1, \dots, 200$



Sobol : $d = 15$, $\{\mathbf{x}_k\}_{14}$ et $\{\mathbf{x}_k\}_{15}$,
 $k = 1, \dots, 200$



3 Critère miniMax & dispersion

3.1 Dispersion (Niederreiter, 1992, Chap. 6)

Discrépance : mesure l'uniformité de la distribution des $\mathbf{x}_k$

⇒ on peut être moins exigeant et s'intéresser seulement au “remplissage” de $\mathcal{X}$ par $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$

$$d_n(\mathbf{X}_n, \mathcal{X}) \triangleq \sup_{\mathbf{x} \in \mathcal{X}} \min_{1 \leq k \leq n} \Delta(\mathbf{x}, \mathbf{x}_k)$$

(▲ on minimise la dispersion ainsi définie ! ▲)

3 Critère miniMax & dispersion

3.1 Dispersion (Niederreiter, 1992, Chap. 6)

Discrépance : mesure l'uniformité de la distribution des $\mathbf{x}_k$

⇒ on peut être moins exigeant et s'intéresser seulement au “remplissage” de $\mathcal{X}$ par $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$

$$d_n(\mathbf{X}_n, \mathcal{X}) \triangleq \sup_{\mathbf{x} \in \mathcal{X}} \min_{1 \leq k \leq n} \Delta(\mathbf{x}, \mathbf{x}_k)$$

(▲ on minimise la dispersion ainsi définie ! ▲)

• $\Delta(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|$ (distance euclidienne)

⇒ $d_n(\mathbf{X}_n, \mathcal{X}) = \phi_{mM}(\mathbf{X}_n) = \text{critère miniMax}$

3 Critère miniMax & dispersion

3.1 Dispersion (Niederreiter, 1992, Chap. 6)

Discrépance : mesure l'uniformité de la distribution des $\mathbf{x}_k$

⇒ on peut être moins exigeant et s'intéresser seulement au “remplissage” de $\mathcal{X}$ par $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$

$$d_n(\mathbf{X}_n, \mathcal{X}) \triangleq \sup_{\mathbf{x} \in \mathcal{X}} \min_{1 \leq k \leq n} \Delta(\mathbf{x}, \mathbf{x}_k)$$

(▲ on minimise la dispersion ainsi définie ! ▲)

- $\Delta(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|$ (distance euclidienne)

$$\Rightarrow d_n(\mathbf{X}_n, \mathcal{X}) = \phi_{mM}(\mathbf{X}_n) = \text{critère miniMax}$$

- $\Delta(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|_\infty = \max_{1 \leq j \leq d} |\{\mathbf{x}\}_j - \{\mathbf{x}'\}_j|$ (distance ℓ_∞)

$$\Rightarrow d_{\infty, n}(\mathbf{X}_n, \mathcal{X}) \text{ (“boules” = cubes, plus faciles à empiler)}$$

3 Critère miniMax & dispersion

3.1 Dispersion (Niederreiter, 1992, Chap. 6)

Discrépance : mesure l'uniformité de la distribution des $\mathbf{x}_k$

⇒ on peut être moins exigeant et s'intéresser seulement au “remplissage” de $\mathcal{X}$ par $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)$

$$d_n(\mathbf{X}_n, \mathcal{X}) \triangleq \sup_{\mathbf{x} \in \mathcal{X}} \min_{1 \leq k \leq n} \Delta(\mathbf{x}, \mathbf{x}_k)$$

(▲ on minimise la dispersion ainsi définie ! ▲)

- $\Delta(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|$ (distance euclidienne)

$$\Rightarrow d_n(\mathbf{X}_n, \mathcal{X}) = \phi_{mM}(\mathbf{X}_n) = \text{critère miniMax}$$

- $\Delta(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|_\infty = \max_{1 \leq j \leq d} |\{\mathbf{x}\}_j - \{\mathbf{x}'\}_j|$ (distance ℓ_∞)

$$\Rightarrow d_{\infty,n}(\mathbf{X}_n, \mathcal{X}) \text{ (“boules” = cubes, plus faciles à empiler)}$$

- $d_{\infty,n}(\mathbf{X}_n, \mathcal{X}) = d_n(\mathbf{X}_n, \mathcal{X})$ pour $d = 1$

- $d_{\infty,n}(\mathbf{X}_n, \mathcal{X}) \leq d_n(\mathbf{X}_n, \mathcal{X}) \leq \sqrt{d} d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$

- $d_{\infty,n}(\mathbf{X}_n, \mathcal{X}) \leq d_n(\mathbf{X}_n, \mathcal{X}) \leq \sqrt{d} d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$
- Recouvrement par des sphères (§ 1.1 et 1.3) :
$$\left(\frac{\text{vol}(\mathcal{X})}{V_d}\right)^{1/d} \frac{1}{n^{1/d}} \leq d_n(\mathbf{X}_n, \mathcal{X}) = \Phi_{mM}(\mathbf{X}_n) \quad (V_d = \text{vol}[\mathcal{B}(0, 1)])$$

- $d_{\infty,n}(\mathbf{X}_n, \mathcal{X}) \leq d_n(\mathbf{X}_n, \mathcal{X}) \leq \sqrt{d} d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$
- Recouvrement par des sphères (§ 1.1 et 1.3) :
$$\left(\frac{\text{vol}(\mathcal{X})}{V_d}\right)^{1/d} \frac{1}{n^{1/d}} \leq d_n(\mathbf{X}_n, \mathcal{X}) = \Phi_{mM}(\mathbf{X}_n) \quad (V_d = \text{vol}[\mathcal{B}(0, 1)])$$
- Recouvrement par des cubes :
$$\frac{1}{2} (\text{vol}(\mathcal{X}))^{1/d} \frac{1}{n^{1/d}} \leq d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$$

- $d_{\infty,n}(\mathbf{X}_n, \mathcal{X}) \leq d_n(\mathbf{X}_n, \mathcal{X}) \leq \sqrt{d} d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$
- Recouvrement par des sphères (§ 1.1 et 1.3) :
$$\left(\frac{\text{vol}(\mathcal{X})}{V_d}\right)^{1/d} \frac{1}{n^{1/d}} \leq d_n(\mathbf{X}_n, \mathcal{X}) = \Phi_{mM}(\mathbf{X}_n) \quad (V_d = \text{vol}[\mathcal{B}(0, 1)])$$
- Recouvrement par des cubes :
$$\frac{1}{2} (\text{vol}(\mathcal{X}))^{1/d} \frac{1}{n^{1/d}} \leq d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$$

Pour $\mathcal{X} = [0, 1]^d$ ($\text{vol}(\mathcal{X}) = 1$)

$$\frac{1}{2} \frac{1}{\lfloor n^{1/d} \rfloor} \leq d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$$

avec égalité pour un certain $\mathbf{X}_n$

- $d_{\infty,n}(\mathbf{X}_n, \mathcal{X}) \leq d_n(\mathbf{X}_n, \mathcal{X}) \leq \sqrt{d} d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$
- Recouvrement par des sphères (§ 1.1 et 1.3) :

$$\left(\frac{\text{vol}(\mathcal{X})}{V_d}\right)^{1/d} \frac{1}{n^{1/d}} \leq d_n(\mathbf{X}_n, \mathcal{X}) = \Phi_{mM}(\mathbf{X}_n) \quad (V_d = \text{vol}[\mathcal{B}(0, 1)])$$
- Recouvrement par des cubes :

$$\frac{1}{2} (\text{vol}(\mathcal{X}))^{1/d} \frac{1}{n^{1/d}} \leq d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$$

Pour $\mathcal{X} = [0, 1]^d$ ($\text{vol}(\mathcal{X}) = 1$)

$$\frac{1}{2} \frac{1}{\lfloor n^{1/d} \rfloor} \leq d_{\infty,n}(\mathbf{X}_n, \mathcal{X}) \leq 2[D_n^*(\mathbf{X}_n)]^{1/d}$$

avec égalité pour un certain $\mathbf{X}_n$

$$\leq A \frac{(\log n)^{(d-1)/d}}{n^{1/d}} \text{ pour } \mathbf{X}_n \text{ à faible discr pance}$$

► faible discr pance $\Rightarrow$ faible dispersion

- $d_{\infty,n}(\mathbf{X}_n, \mathcal{X}) \leq d_n(\mathbf{X}_n, \mathcal{X}) \leq \sqrt{d} d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$
- Recouvrement par des sphères (§ 1.1 et 1.3) :

$$\left(\frac{\text{vol}(\mathcal{X})}{V_d} \right)^{1/d} \frac{1}{n^{1/d}} \leq d_n(\mathbf{X}_n, \mathcal{X}) = \Phi_{mM}(\mathbf{X}_n) \quad (V_d = \text{vol}[\mathcal{B}(0, 1)])$$
- Recouvrement par des cubes :

$$\frac{1}{2} (\text{vol}(\mathcal{X}))^{1/d} \frac{1}{n^{1/d}} \leq d_{\infty,n}(\mathbf{X}_n, \mathcal{X})$$

Pour $\mathcal{X} = [0, 1]^d$ ($\text{vol}(\mathcal{X}) = 1$)

$$\frac{1}{2} \frac{1}{\lfloor n^{1/d} \rfloor} \leq d_{\infty,n}(\mathbf{X}_n, \mathcal{X}) \leq 2[D_n^*(\mathbf{X}_n)]^{1/d}$$

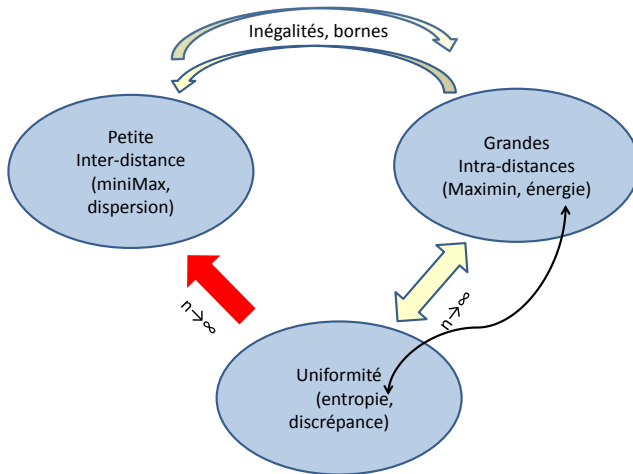
avec égalité pour un certain $\mathbf{X}_n$

$$\leq A \frac{(\log n)^{(d-1)/d}}{n^{1/d}} \text{ pour } \mathbf{X}_n \text{ à faible discr pance}$$

► faible discr pance $\Rightarrow$ faible dispersion

$$\Rightarrow \max \left\{ \frac{1}{(nV_d)^{1/d}}, \frac{1}{2} \frac{1}{\lfloor n^{1/d} \rfloor} \right\} \leq \underbrace{\Phi_{mM}^*}_{\min_{\mathbf{X}_n} \Phi_{mM}(\mathbf{X}_n)} \leq \frac{\sqrt{d}}{2} \frac{1}{\lfloor n^{1/d} \rfloor}$$

(qui am liore un peu les bornes du § 1.2)



3.2 Suites à faible dispersion (Niederreiter, 1992, Chap. 6)

$d = 1$: pour toute suite $\mathbf{X}_\infty$, $\limsup_{n \rightarrow \infty} nd_n(\mathbf{X}_\infty) \geq \frac{1}{2 \log 2} \simeq 0.7213$

3.2 Suites à faible dispersion (Niederreiter, 1992, Chap. 6)

$d = 1$: pour toute suite $\mathbf{X}_\infty$, $\limsup_{n \rightarrow \infty} nd_n(\mathbf{X}_\infty) \geq \frac{1}{2 \log 2} \simeq 0.7213$

La borne est atteinte pour la suite de Ruzsa :

$$x_1 = 1, x_k = \left\{ \frac{\log(2k-3)}{\log 2} \right\}, k \geq 2$$

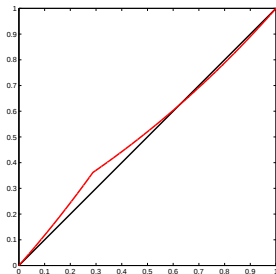
3.2 Suites à faible dispersion (Niederreiter, 1992, Chap. 6)

$d = 1$: pour toute suite $\mathbf{X}_\infty$, $\limsup_{n \rightarrow \infty} nd_n(\mathbf{X}_\infty) \geq \frac{1}{2 \log 2} \simeq 0.7213$

La borne est atteinte pour la suite de Ruzsa :

$$x_1 = 1, x_k = \left\{ \frac{\log(2k-3)}{\log 2} \right\}, k \geq 2$$

mais $D_n^*(\mathbf{X}_\infty) \not\rightarrow 0$! F_n pour $n = 10\,000$ points ($D_n^* \simeq 0.0740$)



Faible discr pance ($\Leftrightarrow$ uniformit ) $\Rightarrow$ faible dispersion
 mais faible dispersion $\not\Rightarrow$ faible discr pance

$$d > 1 : \frac{1}{2} \leq \boxed{\inf_{\mathbf{X}_\infty} \limsup_{n \rightarrow \infty} n^{1/d} d_{\infty,n}(\mathbf{X}_\infty)} \leq \frac{1}{2 \log 2}$$

On connaît des suites qui atteignent la borne supérieure $\frac{1}{2 \log 2}$

On ne connaît pas la valeur de $\inf_{\mathbf{X}_\infty} \limsup_{n \rightarrow \infty} n^{1/d} d_{\infty,n}(\mathbf{X}_\infty)$
(ni a fortiori les meilleures suites $\mathbf{X}_\infty$)

On ne connaît peu de choses concernant $n^{1/d} d_n(\mathbf{X}_\infty) = n^{1/d} \Phi_{mM}(\mathbf{X}_\infty)$!

$$\underline{d} > 1 : \frac{1}{2} \leq \boxed{\inf_{\mathbf{X}_\infty} \limsup_{n \rightarrow \infty} n^{1/d} d_{\infty,n}(\mathbf{X}_\infty)} \leq \frac{1}{2 \log 2}$$

On connaît des suites qui atteignent la borne supérieure $\frac{1}{2 \log 2}$

On ne connaît pas la valeur de $\inf_{\mathbf{X}_\infty} \limsup_{n \rightarrow \infty} n^{1/d} d_{\infty,n}(\mathbf{X}_\infty)$
(ni a fortiori les meilleures suites $\mathbf{X}_\infty$)

On ne connaît peu de choses concernant $n^{1/d} d_n(\mathbf{X}_\infty) = n^{1/d} \Phi_{mM}(\mathbf{X}_\infty)$!

Bornes supérieures (assez pessimistes) :

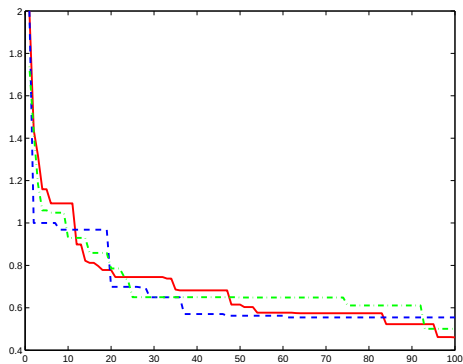
➤ Halton en bases $(b_1, \dots, b_d)$

$$\Rightarrow \boxed{d_{\infty,n}(\mathbf{X}_\infty) < \frac{\max_{1 \leq i \leq d} b_i}{n^{1/d}}}$$

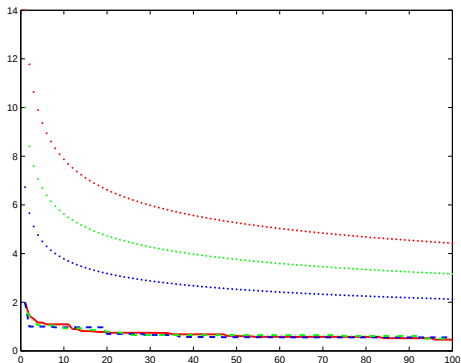
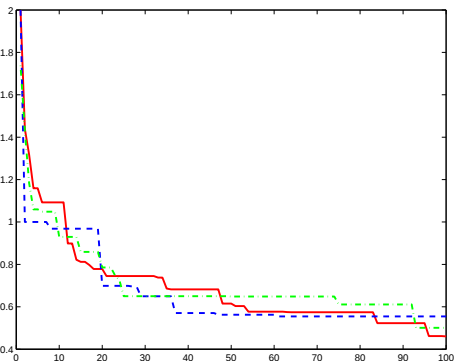
➤ suite- (t, d) en base b

$$\Rightarrow \boxed{d_{\infty,n}(\mathbf{X}_\infty) < \frac{b^{1+t/d}}{n^{1/d}}}$$

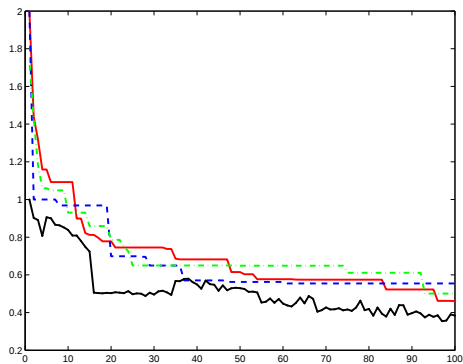
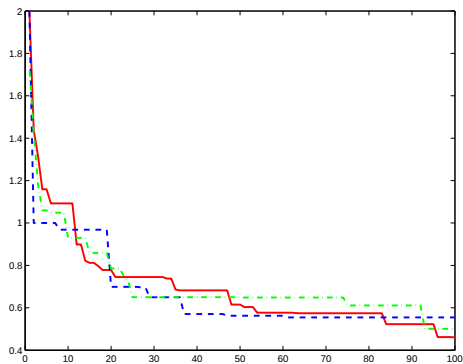
Ex. : $d = 4$, n de 1 à 100, $\Phi_{mM}(\mathbf{X}_n)$ pour **Halton** ($\mathbf{b} = (2, 3, 5, 7)$), **Sobol** ($t = 3$, $b = 2$) et **Faure** ($t = 0$, $b = 5$)



Ex. : $d = 4$, n de 1 à 100, $\Phi_{mM}(\mathbf{X}_n)$ pour **Halton** ($\mathbf{b} = (2, 3, 5, 7)$), **Sobol** ($t = 3$, $b = 2$) et **Faure** ($t = 0$, $b = 5$)



Ex. : $d = 4$, n de 1 à 100, $\Phi_{mM}(\mathbf{X}_n)$ pour **Halton** ($\mathbf{b} = (2, 3, 5, 7)$), **Sobol** ($t = 3$, $b = 2$) et **Faure** ($t = 0$, $b = 5$)



avec algorithme kmeans sur 10 000 points (suite de Sobol), initialisation des centres aux n premiers points

4 Conclusions partie (1) — sans modèle

- Beaucoup de critères disponibles (géométrie, uniformité)
 $\Phi_{mM}(\cdot)$ intéressant, mais calculable seulement pour d petit ($d \lesssim 5$)

4 Conclusions partie (1) — sans modèle

- Beaucoup de critères disponibles (géométrie, uniformité)
 $\Phi_{mM}(\cdot)$ intéressant, mais calculable seulement pour d petit ($d \lesssim 5$)
- La plupart sont difficiles à optimiser (non convexes, multimodaux, parfois non différentiables) $\Rightarrow$ seulement pour d petit

4 Conclusions partie (1) — sans modèle

- Beaucoup de critères disponibles (géométrie, uniformité)
 $\Phi_{mM}(\cdot)$ intéressant, mais calculable seulement pour d petit ($d \lesssim 5$)
- La plupart sont difficiles à optimiser (non convexes, multimodaux, parfois non différentiables) $\Rightarrow$ seulement pour d petit
- Suites à faible discrédance :
 - plans peu coûteux à générer
 - points $\mathbf{x}_k$ bien mélangés (pas nécessaire de fixer n *a priori*)
 - s'adapte à $\mathcal{X}$ compact quelconque (d'intérieur non vide)
 - (générer dans un cube contenant $\mathcal{X}$, puis rejeter les points hors de $\mathcal{X}$)

4 Conclusions partie (1) — sans modèle

- Beaucoup de critères disponibles (géométrie, uniformité)
 $\Phi_{mM}(\cdot)$ intéressant, mais calculable seulement pour d petit ($d \lesssim 5$)
- La plupart sont difficiles à optimiser (non convexes, multimodaux, parfois non différentiables) $\Rightarrow$ seulement pour d petit
- Suites à faible discrédance :
 plans peu coûteux à générer
 points $\mathbf{x}_k$ bien mélangés (pas nécessaire de fixer n *a priori*)
 s'adapte à $\mathcal{X}$ compact quelconque (d'intérieur non vide)
 (générer dans un cube contenant $\mathcal{X}$, puis rejeter
 les points hors de $\mathcal{X}$)

Une méthode efficace et peu coûteuse (pour $\mathcal{X}$ compact convexe) :

4 Conclusions partie (1) — sans modèle

- Beaucoup de critères disponibles (géométrie, uniformité)
 $\Phi_{mM}(\cdot)$ intéressant, mais calculable seulement pour d petit ($d \lesssim 5$)
- La plupart sont difficiles à optimiser (non convexes, multimodaux, parfois non différentiables) $\Rightarrow$ seulement pour d petit
- Suites à faible discrédance :
 plans peu coûteux à générer
 points $\mathbf{x}_k$ bien mélangés (pas nécessaire de fixer n *a priori*)
 s'adapte à $\mathcal{X}$ compact quelconque (d'intérieur non vide)
 (générer dans un cube contenant $\mathcal{X}$, puis rejeter
 les points hors de $\mathcal{X}$)

Une méthode efficace et peu coûteuse (pour $\mathcal{X}$ compact convexe) :

- générer N points dans $\mathcal{X}$ avec une suite à faible discrédance (N grand)

4 Conclusions partie (1) — sans modèle

- Beaucoup de critères disponibles (géométrie, uniformité)
 $\Phi_{mM}(\cdot)$ intéressant, mais calculable seulement pour d petit ($d \lesssim 5$)
- La plupart sont difficiles à optimiser (non convexes, multimodaux, parfois non différentiables) $\Rightarrow$ seulement pour d petit
- Suites à faible discrédance :
 plans peu coûteux à générer
 points $\mathbf{x}_k$ bien mélangés (pas nécessaire de fixer n *a priori*)
 s'adapte à $\mathcal{X}$ compact quelconque (d'intérieur non vide)
 (générer dans un cube contenant $\mathcal{X}$, puis rejeter les points hors de $\mathcal{X}$)

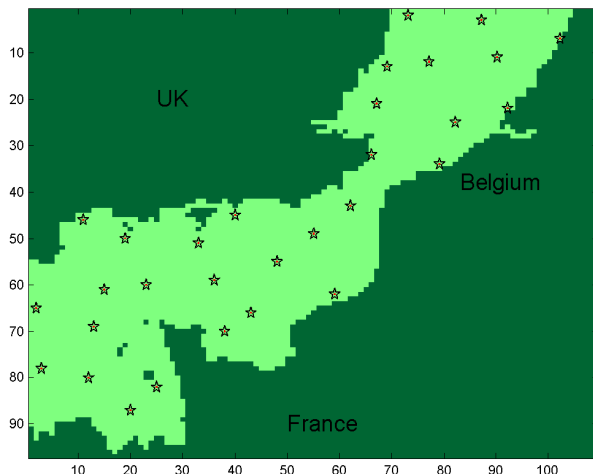
Une méthode efficace et peu coûteuse (pour $\mathcal{X}$ compact convexe) :

- générer N points dans $\mathcal{X}$ avec une suite à faible discrédance (N grand)
- utiliser un algorithme de type `kmeans` pour former n amas (*clusters*)

Ex. : $\mathbf{X}_n = 30$ points dans un tronçon de la mer du Nord

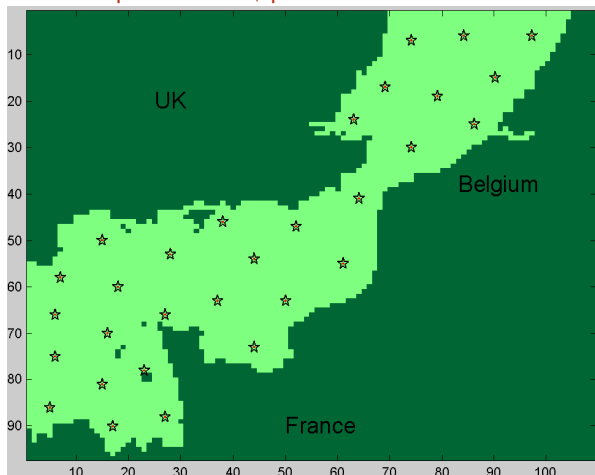
Ex. : $\mathbf{X}_n = 30$ points dans un tronçon de la mer du Nord

30 points Sobol



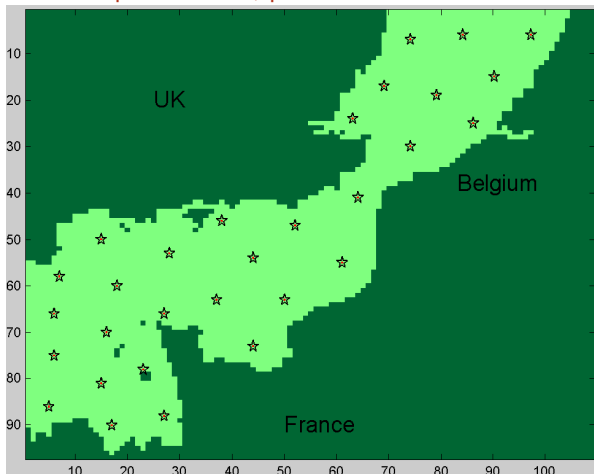
Ex. : $\mathbf{X}_n = 30$ points dans un tronçon de la mer du Nord

1 000 points Sobol, puis kmeans avec 30 amas



Ex. : $\mathbf{X}_n = 30$ points dans un tronçon de la mer du Nord

1 000 points Sobol, puis kmeans avec 30 amas



On peut toujours ensuite optimiser $\mathbf{X}_n$,
rajouter de nouveaux points. . .

Mais le fléau de la dimension est toujours là !

$d = 50$, suite-(0, d) de Faure $\implies b \geq d$ premier $\rightarrow b = 53$

$n = b^q$ pour garantir un point dans chaque pavé coupé suivant q dimensions

$q = 2 \implies n = 2809$

$q = 50 \implies n \simeq 1.6360 \cdot 10^{86}$

Mais le fléau de la dimension est toujours là !

$d = 50$, suite-(0, d) de Faure $\implies b \geq d$ premier $\rightarrow b = 53$

$n = b^q$ pour garantir un point dans chaque pavé coupé suivant q dimensions

$$q = 2 \implies n = 2809$$

$$q = 50 \implies n \simeq 1.6360 \cdot 10^{86}$$

$d = 50$, suite-(t , d) en base 2, avec t le + petit possible $\rightarrow t = 77$

Pour que chaque pavé élémentaire coupé suivant chaque dimension contienne des points, il faut $n = b^m$ avec $\sum_{j=1}^d q_j = d = m - t \implies m = 127$ et $n \simeq 1.7014 \cdot 10^{38}$

Mais le fléau de la dimension est toujours là !

$d = 50$, suite-(0, d) de Faure $\Rightarrow b \geq d$ premier $\rightarrow b = 53$

$n = b^q$ pour garantir un point dans chaque pavé coupé suivant q dimensions

$q = 2 \Rightarrow n = 2809$

$q = 50 \Rightarrow n \simeq 1.6360 \cdot 10^{86}$

$d = 50$, suite-(t, d) en base 2, avec t le + petit possible $\rightarrow t = 77$

Pour que chaque pavé élémentaire coupé suivant chaque dimension contienne des points, il faut $n = b^m$ avec $\sum_{j=1}^d q_j = d = m - t \Rightarrow m = 127$ et $n \simeq 1.7014 \cdot 10^{38}$

Voir (Owen, 1998) pour des constructions possibles, du type

$$\mathbf{X}_n = \{\mathbf{X}_n\}_{1:d} = \begin{pmatrix} \{\mathbf{X}_n\}_{1:s} \\ \{\mathbf{X}_n\}_{s+1:d} \end{pmatrix}$$

avec par ex. $\{\mathbf{X}_n\}_{1:s}$ à faible discrédance

$\{\mathbf{X}_n\}_{s+1:d}$ un hypercube latin

Références 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/Schneberg (Austria)*. Physica Verlag, Heidelberg, pp. 13–20.
- Botkin, N., Turova-Botkina, V., 1994. An algorithm for finding the Chebyshev center of a convex polyhedron. *Applied Mathematics and Optimization* 29, 211–222.
- Clarke, F., 1975. Generalized gradients and applications. *Transactions of the American Mathematical Society* 205, 247–262.
- 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.
- Dem'yanov, V., 1968. Algorithms for some minimax problems. *Journal of Computer and System Sciences* 2 (4), 342–380.
- Dem'yanov, V., Malozemov, V., 1974. *Introduction to Minimax*. Dover, New York.

Références 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. Discrepances 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.
- Hardin, D., Saff, E., 2004. Discretizing manifolds via minimum energy points. *Notices of the AMS* 51 (10), 1186–1194.
- 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: Heilekalek, P., Larcher, G. (Eds.), *Random and quasi-random point sets*. Vol. 138 of *Lecture Notes in Statist.* Springer, New York, pp. 109–166.

Références III

- Hlawka, E., 1961. Funktionen von beschränkter variation in der theorie der gleichverteilung. *Annali di Matematica Pura ed Applicata* 54 (1), 325–333.
- John, F., 1985. Extremum problems with inequalities as subsidiary conditions. In: Moser, J. (Ed.), *F. John Collected Papers*, Vol. 2. Birkhäuser, Boston, pp. 543–560, originally published in *Courant Anniversary Volume, 1948*, pp. 187–204.
- Johnson, M., Moore, L., Ylvisaker, D., 1990. Minimax and maximin distance designs. *Journal of Statistical Planning and Inference* 26, 131–148.
- Joseph, V., Dasgupta, T., Tuo, R., Wu, C., 2014. Sequential exploration of complex surfaces using minimum energy designs. *Technometrics*(to appear).
- 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.
- Leonenko, N., Pronzato, L., Savani, V., 2008. A class of Rényi information estimators for multidimensional densities. *Annals of Statistics* 36 (5), 2153–2182 (Correction in AS, 38(6):3837–3838, 2010).
- 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.

Références IV

- Lubachevsky, B., Stillinger, F., 1990. Geometric properties of random disk packings. *Journal of Statistical Physics* 60 (5-6), 561–583.
- 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.
- 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), 239–245.
- Morris, M., Mitchell, T., 1995. Exploratory designs for computational experiments. *Journal of Statistical Planning and Inference* 43, 381–402.
- 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.
- 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.

Références V

- Persson, P.-O., Strang, G., 2004. A simple mesh generator in MATLAB. *SIAM review* 46 (2), 329–345.
- Pronzato, L., Müller, W., 2012. Design of computer experiments: space filling and beyond. *Statistics and Computing* 22, 681–701, DOI 10.1007/s11222-011-9242-3.
- 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.
- 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.
- 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.
- van der Corput, J., 1935. Verteilungsfunktionen. I. *Mitt. Proc. Akad. Wet. Amsterdam* 38, 813–821.

Références VI

- 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.
- 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. <https://hal.archives-ouvertes.fr/hal-00955294>.
- Welzl, E., 1991. Smallest enclosing disks (balls and ellipsoids). In: *Lecture Notes in Computer Science*, vol. 555. Springer, Heidelberg, pp. 359–370.
- Xu, S., Freund, R., Sun, J., 2003. Solution methodologies for the smallest enclosing circle problem. *Computational Optimization and Applications* 25 (1-3), 283–292.
- Yukich, J., 1998. *Probability Theory of Classical Euclidean Optimization Problems*. Springer, Berlin.