

Workshop



Plans d'expériences

10 -12 Septembre 2013

Hostellerie La Magnaneraie
VILLENEUVE LÈS AVIGNON

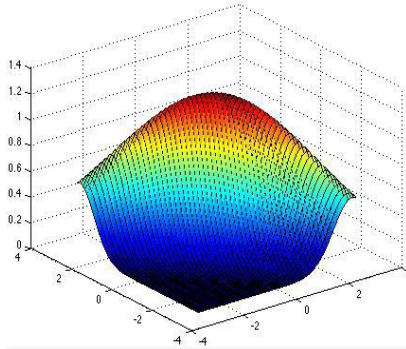
Plans uniformes

Domaine avec contraintes

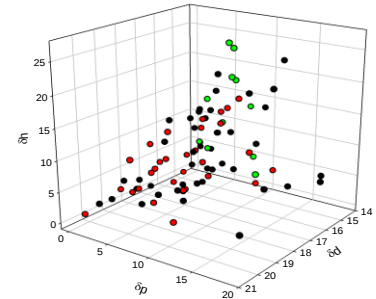
Plans d'expériences



*Expériences
numériques*

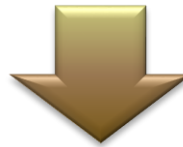


*Surfaces de réponse
complexes*



*QSAR/QSPR
Points candidats*

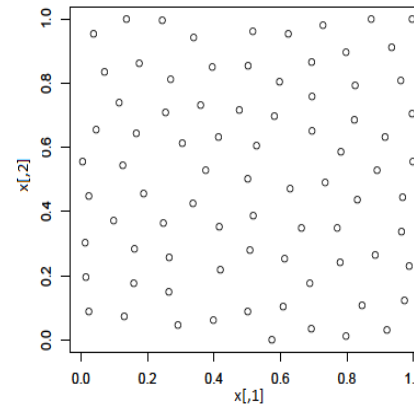
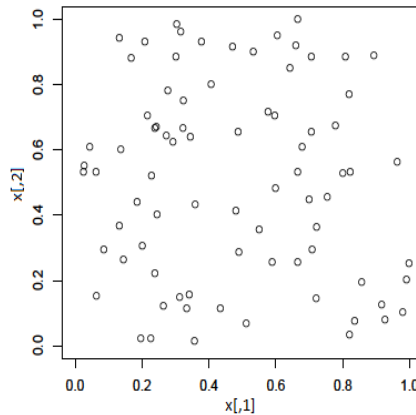
- ➔ Grand nombre de paramètres ($k > 20$)
- ➔ Large domaine de variation
- ➔ Phénomène complexe : Surface de réponse chaotique
- ➔ Echantillonnage, choix dans un ensemble candidat, ...



Plans d'expériences spécifiques

SFD : Space Filling Design

Un plan d'expériences uniforme (Space-Filling Design, SFD) a la particularité de répartir les points expérimentaux uniformément dans le domaine de variation des variables d'entrée



Nécessité de définir des critères intrinsèques de qualité en terme d'uniformité

Les plans uniformes

⇒ Critères pour caractériser l'uniformité d'un plan :

→ La **discrépance** : mesure l'**écart** entre une distribution empirique de points donnée et une distribution théorique de points uniforme ↘

→ La **distance minimale**, Mindist : **valeur minimale des distances minimales entre 2 points** ↗

$$MinDist = \min_{x_i \in X} \min_{\substack{x_k \in X \\ k \neq i}} dist(x^i, x^k)$$

→ La **moyenne des distances minimales**, Moydist : **valeur moyenne des distances minimales**

→ La **mesure de recouvrement**, Cov : **coefficient de variation des distances minimales** ↘

$$Cov = \frac{1}{\gamma} \frac{1}{n} \left(\sum_{i=1}^n (\gamma_i - \bar{\gamma})^2 \right)^{1/2}$$

$$\gamma_i = \min_{k \neq i} dist(x^i, x^k) \quad \bar{\gamma} = \frac{1}{n} \sum_{i=1}^n \gamma_i$$

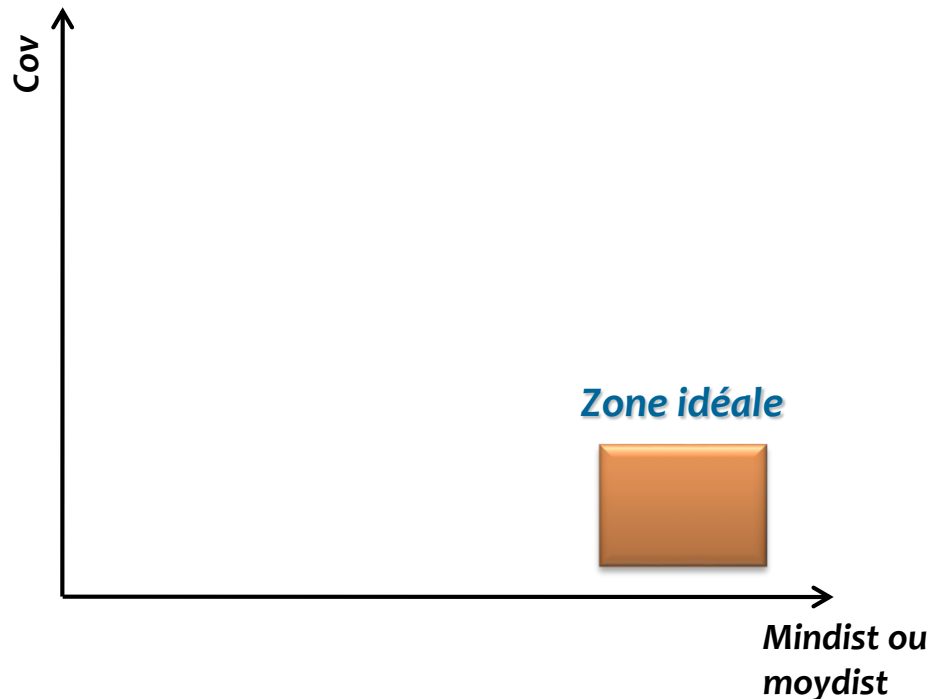


Les plans uniformes

⇒ Critères pour caractériser l'uniformité d'un plan :



- **Maximiser** la valeur du **Mindist** et du **Moydist** pour ne pas avoir d'amas ,
et
- **Minimiser** la valeur **Cov** pour avoir une répartition homogène.

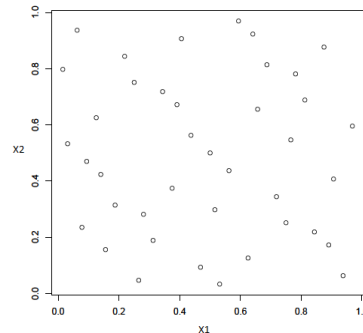


Les plans uniformes

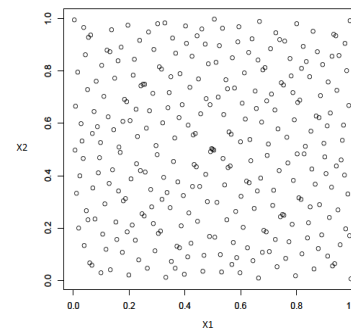
Quelques plans uniformes

↳ *basés sur des lois de distribution des points :*

- **Suites à faible discrédance :** *suites de Halton, séquences de Hammersley, suites de Faure, suites de Sobol*

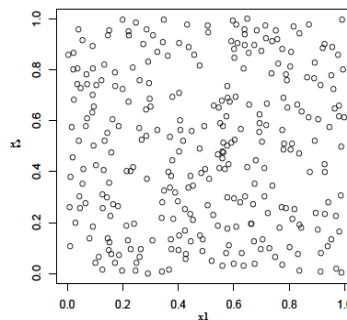


suite de Faure ($N = 40$)



suite de Sobol ($N = 300$)

- **Plans à entropie maximale :** *Plans Dmax*



($N = 300$)



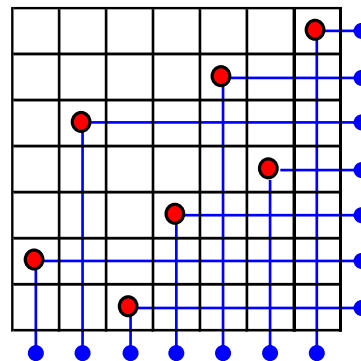
Les plans uniformes

Quelques plans uniformes

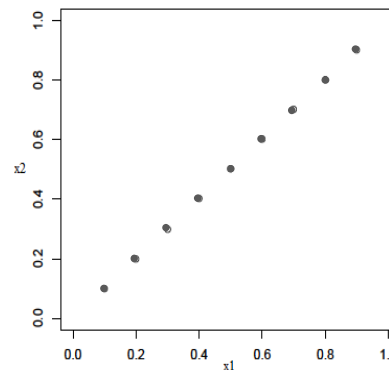
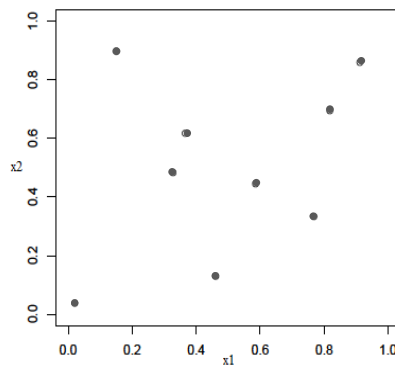
↳ basés sur des contraintes marginales :

- **Hypercubes Latins** : Random LHS, Maximin LHS, Improved LHS

x_1	x_2
1	2
2	5
3	1
4	3
5	6
6	4
7	7



projections sur les axes des variables les plus uniformes possibles



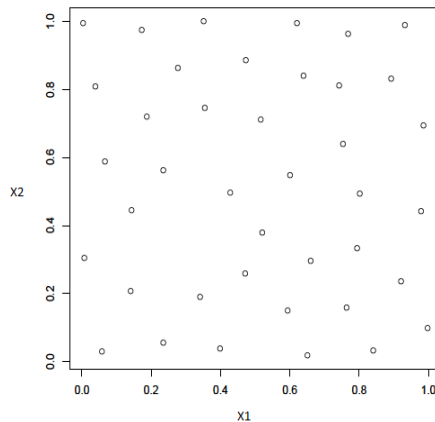
(N = 9)

Les plans uniformes

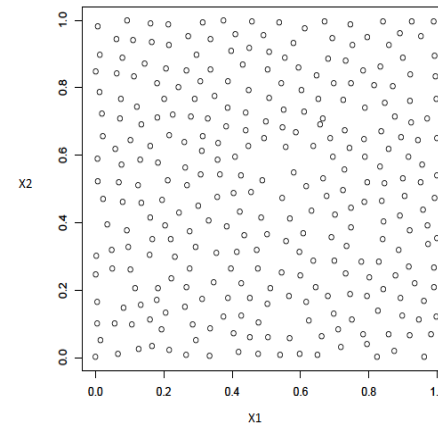
Quelques plans uniformes

↳ basés sur le principe de répulsion entre points :

- **Plans de Strauss**: inspiré du principe de répulsion entre particules



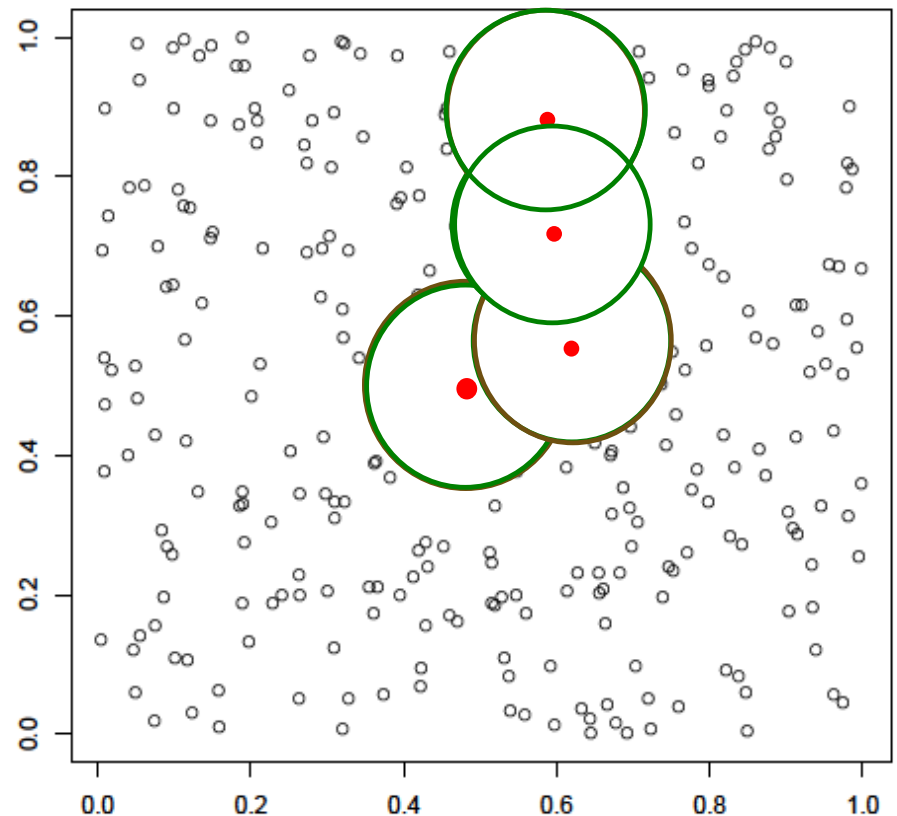
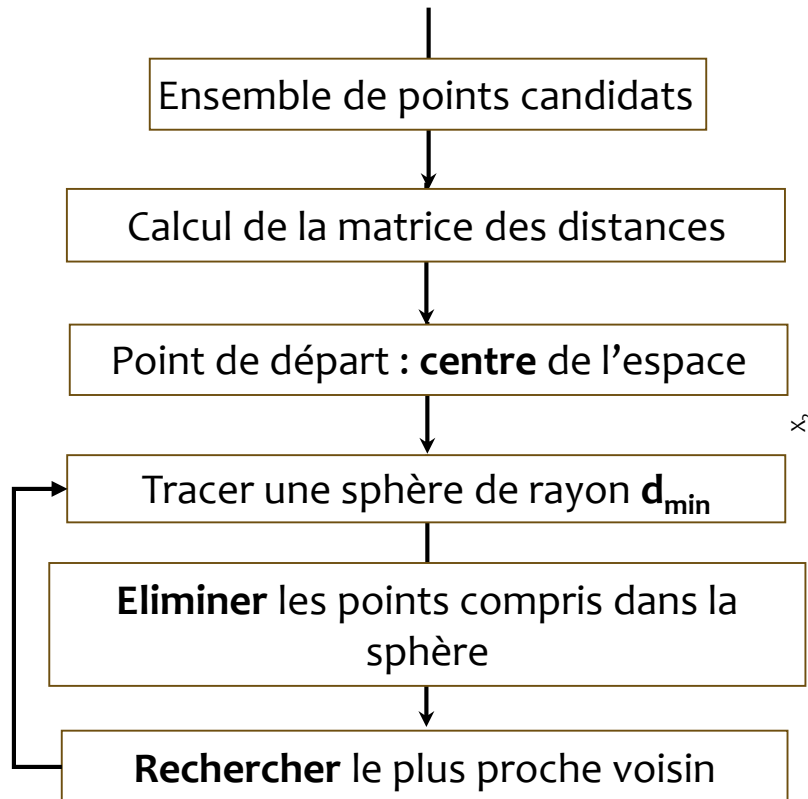
($N = 40$)



($N = 300$)

Les plans uniformes

➡ Algorithme WSP : algorithme de construction de plans uniformes



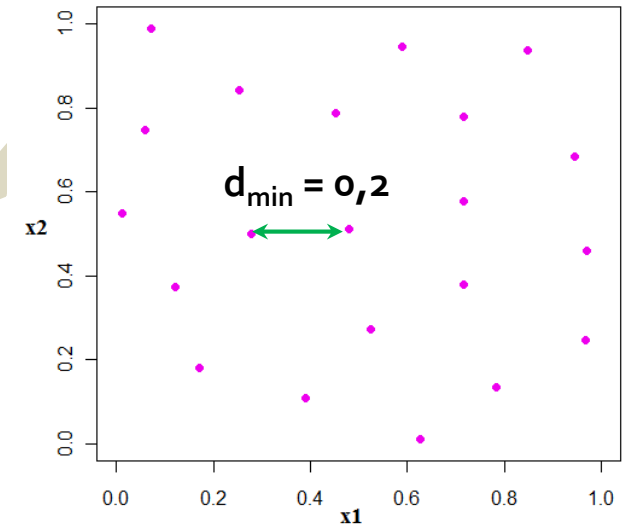
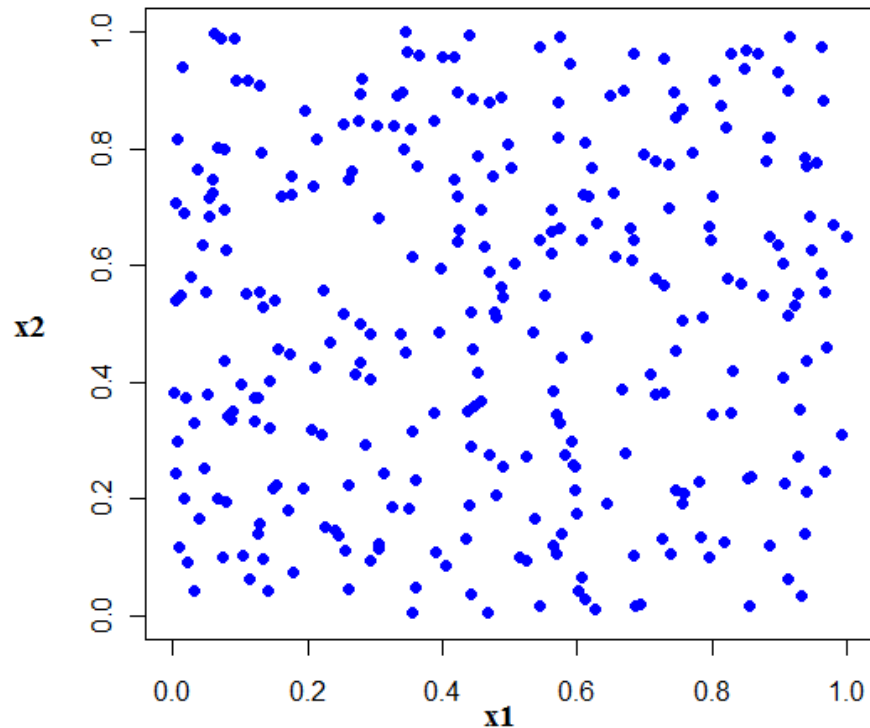
Santiago, J., Claeys-Bruno, M., Sergent, M. 2012. Construction of space filling designs using WSP algorithm for high dimensional spaces. *Chemometrics and Intelligent Laboratory System* 113 :26-31.

Les plans uniformes

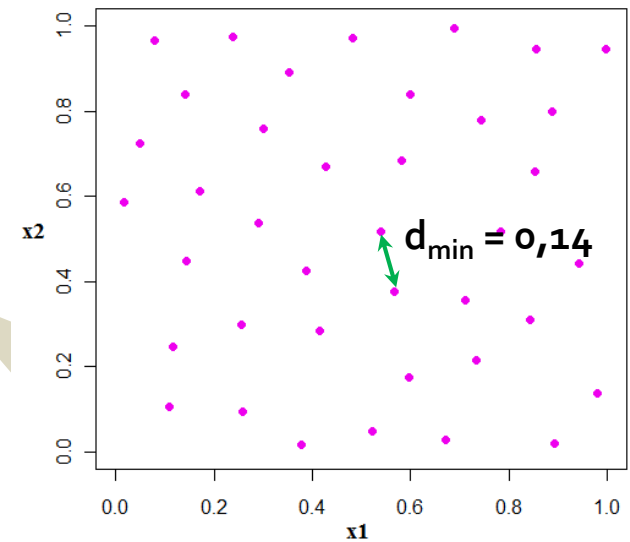
➡ Algorithme WSP :

ensemble de points candidats

N = 300

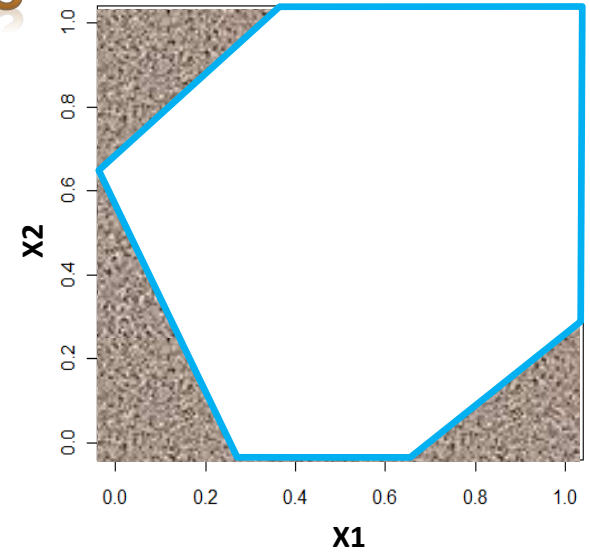


N=40 points

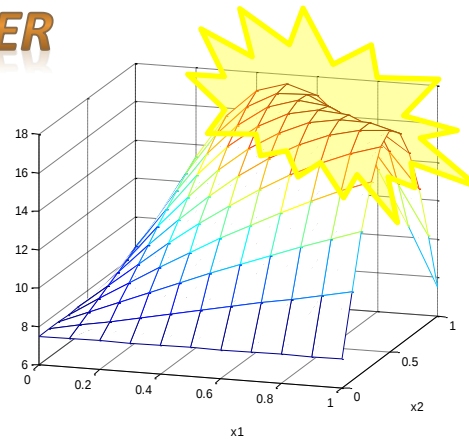


Les plans uniformes

→ **DOMAINE EXPÉRIMENTAL AVEC CONTRAINTES**



→ **ZONE À DENSIFIER**



Les plans uniformes

→ Algorithme WSP :



WSP évolutif (aWSP)

*Pour augmenter la densité de doints, la distance $d_{\min}$
entre deux points sera fonction de la position du
point considéré dans l'espace des variables*

Algorithme α WSP

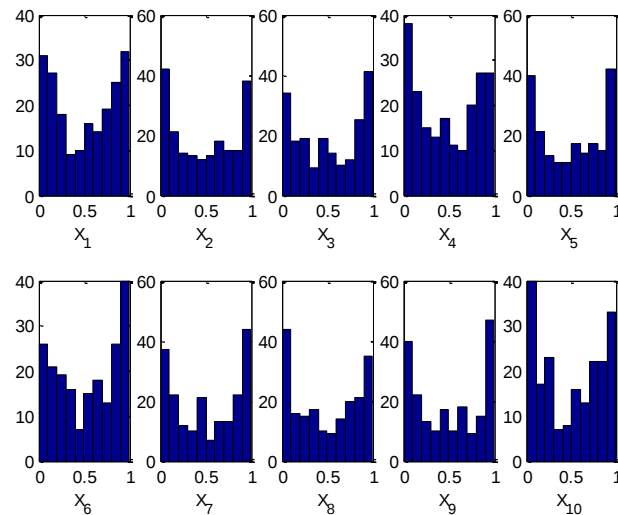
→ **ZONE À DENSIFIER** ✱ Au centre du domaine "centre creux" :

En grande dimension, les SFD subissent le fléau de la grande dimension à savoir une **zone centrale "creuse"**.

Il a été montré que dans une distribution de points en grande dimension, la probabilité qu'ils se situent dans les coins de l'hypercube tend vers 1 quand la dimension augmente.



Effectifs sur chaque axe factoriel



plan WSP à 200 points en 10D

Algorithme aWSP

→ ZONE À DENSIFIER ✖ Au centre du domaine :

- La valeur d_{min} est calculée à chaque itération selon une relation non linéaire, fonction de la position du point :

$$d_{min} = \left[\frac{\text{distance entre le point considéré et le centre} - \text{distance minimale}}{\text{distance maximale} - \text{distance minimale}} \right]^r$$

distance minimale = $\min_{x_i \in X} \text{dist}(x_i, x_{\text{centre}})$

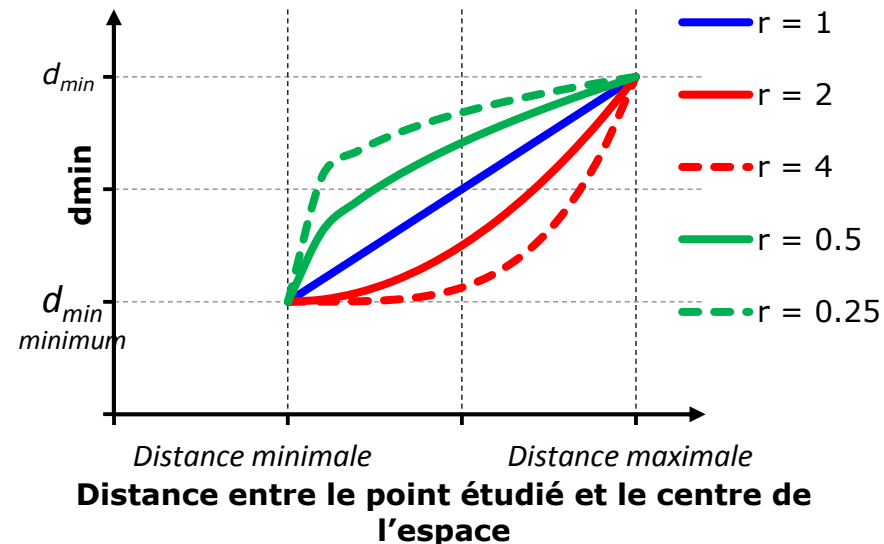
distance maximale = $\max_{x_i \in X} \text{dist}(x_i, x_{\text{centre}})$

- La valeur de l'exposant r influence la courbure du graphe des variations de d_{min} :

- $r = 1$, d_{min} suit une variation linéaire

- $r > 1$, la variation du d_{min} favorise la sélection de points situés à des distances faibles, i.e., d_{min} augmente lentement, permettant la sélection d'un grand nombre de points au centre du domaine

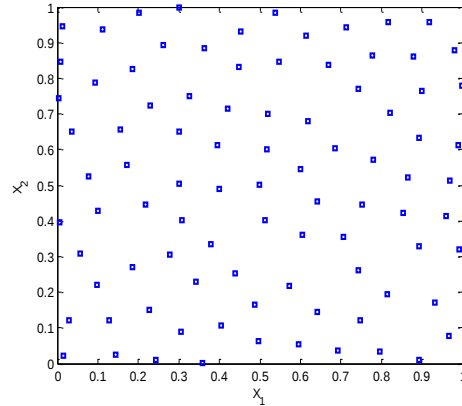
- $r < 1$, la variation faible du d_{min} au centre du domaine permet une augmentation légère de la densité des points au centre du domaine, avec une faible densité à la périphérie



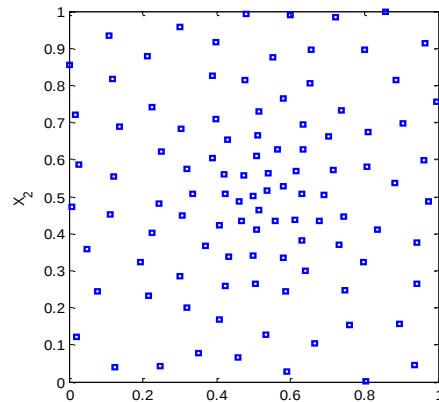
Algorithme α WSP

→ **ZONE À DENSIFIER** Au centre du domaine :

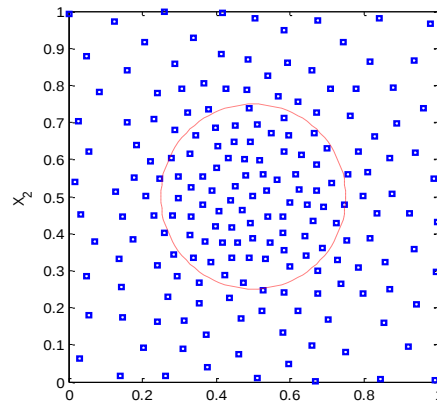
WSP classique permettant la sélection de 90 points pour un $d_{\min} = 0.1$.



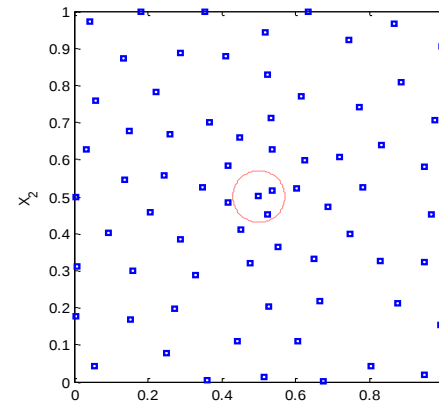
Etude de la variation de la courbure sur un WSP "évolutif" pour un $d_{\min}$ minimum = 0.04 et $d_{\min}$ maximum = 0.15.



$r = 1$
 $N = 105$ points.



$r = 2$
 $N = 193$ points.



$r = 0.5$
 $N = 71$ points

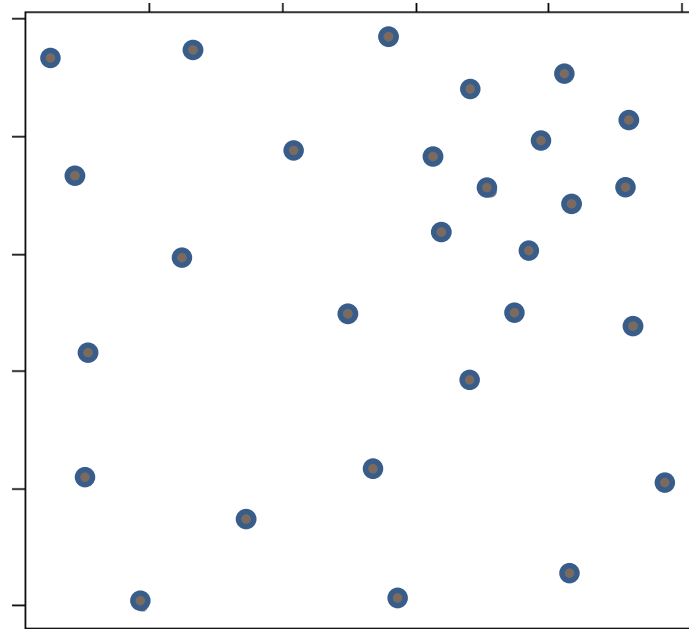
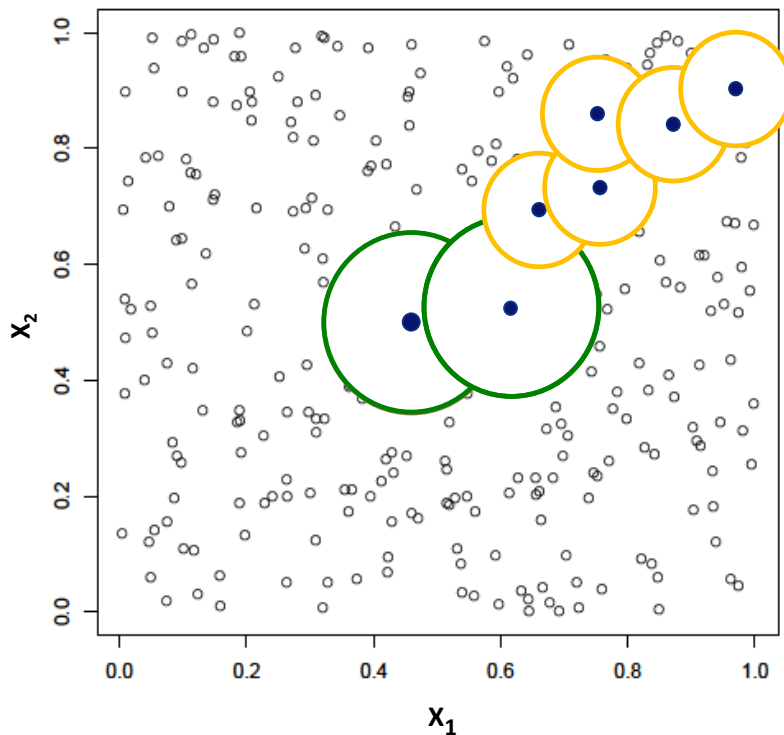
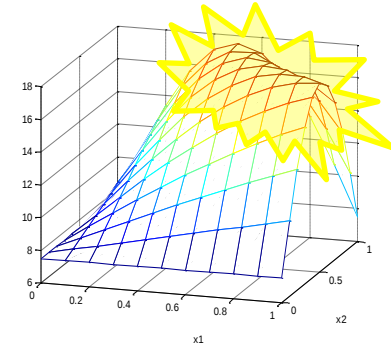


Algorithme α WSP

→ **ZONE À DENSIFIER**

✧ Zones particulières :

➡ **Phénomène complexe : Surface de réponse chaotique**



Algorithme aWSP

→ ZONE À DENSIFIER

• Contraintes individuelles

La zone d'intérêt à densifier est définie par :

$$X_1 \geq 0.6$$

$$X_2 \geq 0.7$$

$$X_3 \leq 0.2$$

Algorithme WSP classique

$$d_{\min} = 0.15$$

$N = 287$ points

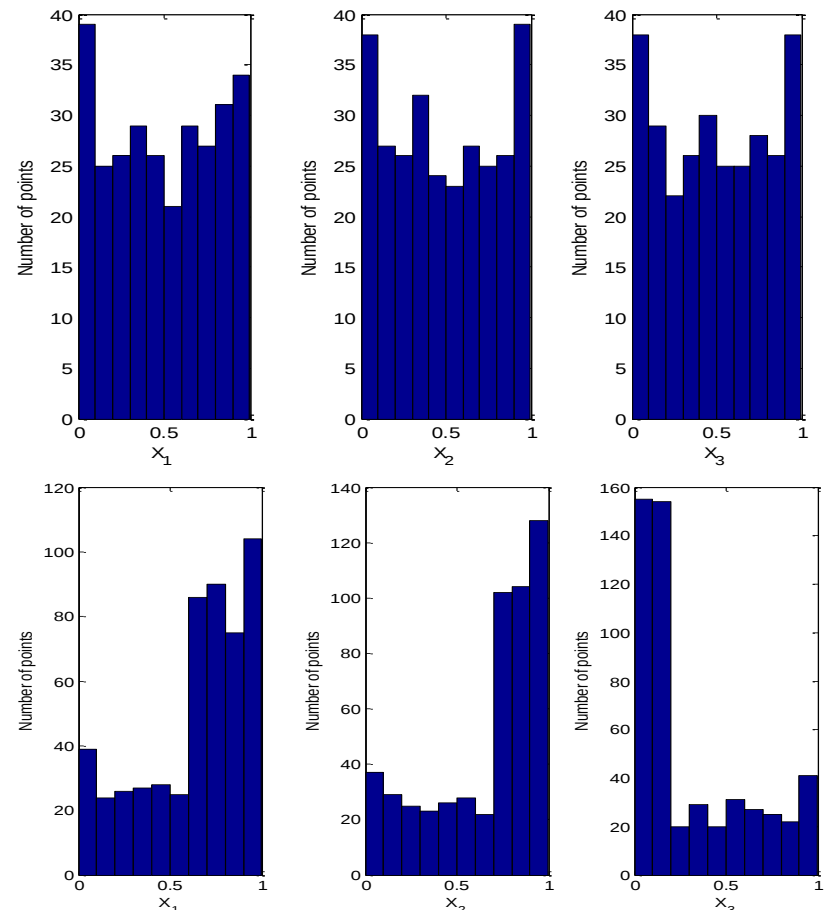
Matrice de points candidats en 3D
et 100 000 points

Algorithme WSP évolutif

$$d_{\min \text{ in zone}} = 0.04$$

$$d_{\min \text{ out zone}} = 0.15$$

$N = 524$ points



Algorithme aWSP

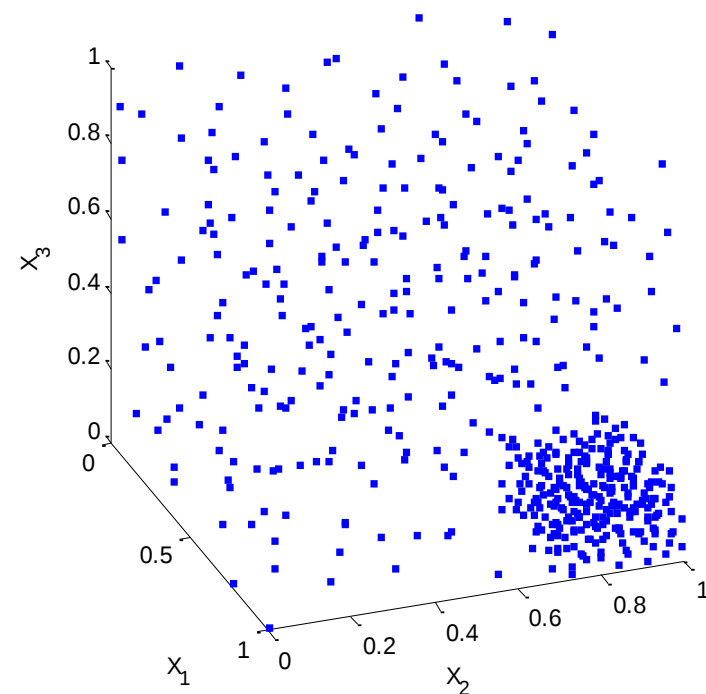
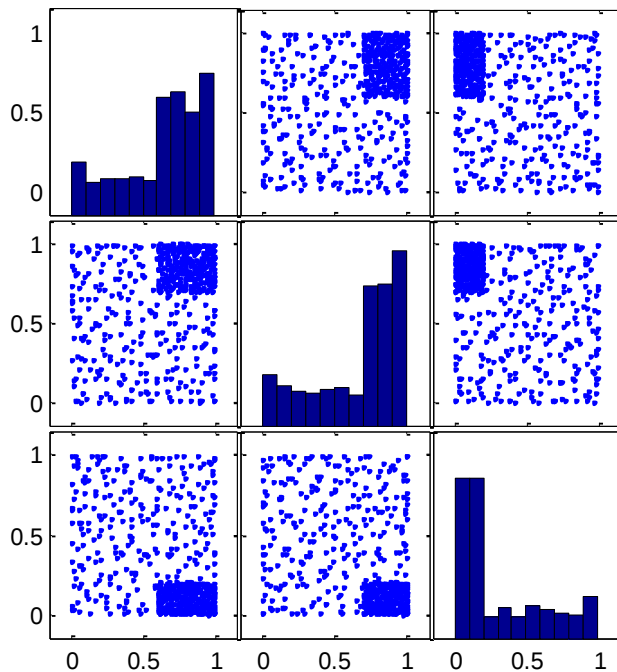
→ **ZONE À DENSIFIER**

La zone d'intérêt à densifier est définie par :

$$X_1 \geq 0.6$$

$$X_2 \geq 0.7$$

$$X_3 \leq 0.2$$



N = 524 points

Algorithme aWSP

• Contraintes relationnelles

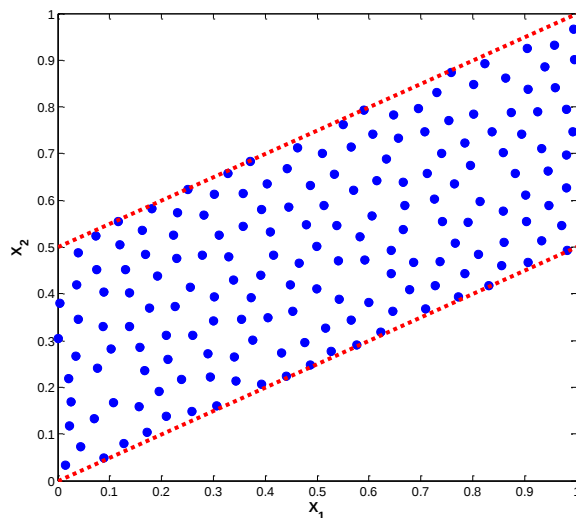
La zone d'intérêt à densifier est définie par : $0.5 X_1 < X_2 < 0.5 X_1 + 0.5$

3 possibilités

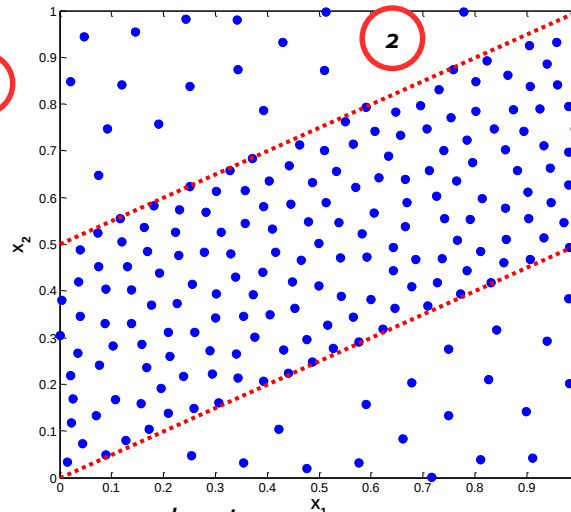
On souhaite étudier
seulement la zone d'intérêt

On souhaite étudier l'ensemble
de l'espace en augmentant la
densité dans la zone d'intérêt

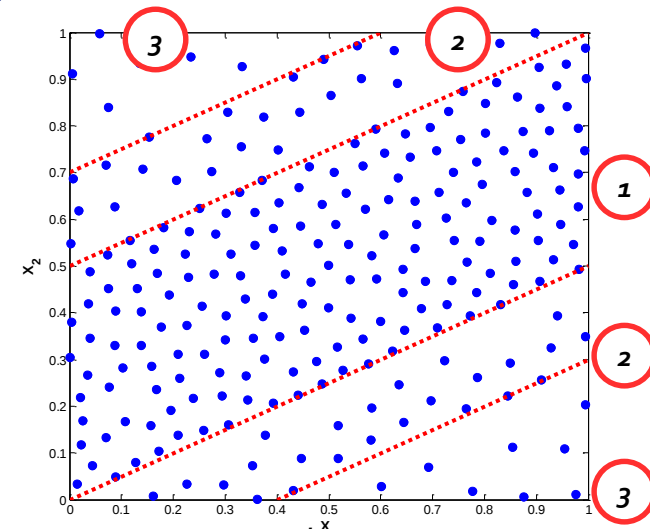
On souhaite étudier
l'ensemble de l'espace avec
des densités variables



d_{min} in zone 1 = 0.05
 $N = 176$ points



d_{min} in zone 1 = 0.05
 d_{min} in zone 2 = 0.1
 $N = 214$ points



d_{min} in zone 1 = 0.05
 d_{min} in zone 2 = 0.07
 d_{min} in zone 3 = 0.1
 $N = 240$ points

