

MA-510-41-2

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique
UNIVERSITE SAAD DAHLEB-BLIDA



FACULTE DES SCIENCES
DEPARTEMENT DE MATHÉMATIQUES

MEMOIRE DE MASTER

En Mathématiques

Spécialité : Modélisation Stochastique et Statistique

Thème

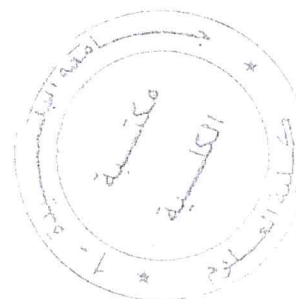
Planification d'expériences numériques par la méthode
de Monte Carlo par chaînes de Markov

Présenté par :

FAKIR YUCEF

Devant le jury composé de :

Président : Mr O. TAMI
Examineurs : Mr F.HANNANE
Promoteur : Mr H. EL MOSSAOUI



2016/2017

MA-510-41-2

ملخص

منهجية البحوث التجريبية هي مجموعة من الأساليب وطرق التفكير لأي مجرب راغب في التخطيط التجريبي. وهو يهدف إلى تمكينها من تحسين فعالية بحوثه التجريبية بغض النظر عن نشاطها. و سوف تساعدنا على التعبير عن مشاكلنا بشكل أفضل واقتراح استراتيجيات التجريبية الأمثل استنادا إلى الأهداف التي وضعتها.

والنقطة الرئيسية في هذا العمل هي لدراسة معايير المثالية مثل انتظام دسكخبونس و انس تخبي والنظر في طريقة جديدة لتصميم التجريبي القائم على عينات المحاكاة وفقا للتوزيعات الاحتمالية على طريقة مونت كارلو بواسطة سلاسل ماركوف.

Résumé

La méthodologie de la recherche expérimentale (MRE) est un ensemble de méthodes et de modes de raisonnement destiné à tout expérimentateur désirant faire de la planification expérimentale. Elle a pour objet de lui permettre d'optimiser l'efficacité de sa recherche expérimentale quel que soit sa branche d'activité. Pour ce faire, elle va l'aider à exprimer aux mieux son problème et lui proposer des stratégies expérimentales optimales en fonction des objectifs qu'il s'est fixé et des moyennes dont il dispose.

Le point essentiel de ce travail consiste à étudier des critères d'optimalité comme celui d'uniformité, de discrèpence et d'entropie et d'envisager une nouvelle méthode de planification d'expériences basée sur la simulation des échantillons selon des lois de probabilité par une Méthode de Monte Carlo par chaînes de Markov (McMC).

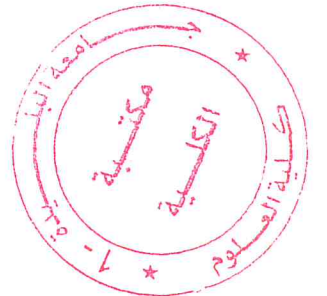


Table des matières

Remerciements	5
Principaux notions et symboles	6
Liste des illustrations, graphiques et tableaux	8
Liste des algorithmes	10
Introduction	11
I METHODOLOGIE DE LA RECHERCHE EXPERIMENTALE	14
1.1 Historique	14
1.2 Intérêts des plans d'expériences	15
1.3 Notion d'espace expérimental	16
1.3.1 Domaine d'étude (DE)	17
1.3.2 Surfaces de réponse	18
1.3.3 Variables centrées réduites	19
1.3.4 Matrice d'expérience	20
1.4 Modélisation mathématique a priori de la réponse [5]	21
1.4.1 Modélisation mathématique	21
1.4.2 Modélisation expérimentale	22
1.4.3 Modèle du statisticien	23

1.4.4	Système d'équations	23
1.4.5	Estimation des coefficients par la méthode des moindres carrés	24
1.4.6	Propriétés des coefficients	26
1.5	Plans classiques	28
1.5.1	Plans de criblages	28
1.5.2	Plans de modélisation	29
1.5.3	Plans de mélanges	29
II	CRITÈRES ET PLANS OPTIMAUX	30
2.1	Critères d'optimalités classique	30
2.1.1	Qualité de la représentation des coefficients	31
2.1.2	Critère A	32
2.1.3	Critère D	32
2.1.4	Critère E	33
2.1.5	Critère G	33
2.1.6	Critère M	34
2.1.7	Critère-R	34
2.1.8	Critère d'orthogonalité	35
2.1.9	Critère de Presque Orthogonalité	35
2.1.10	Critère d'isovariation par rotation	36

2.2	critère d'uniformité basé sur la discrédance	36
2.2.1	La discrédance	36
2.2.2	Discrédance en norme 2	38
2.2.3	Exemples	39
2.3	Critères de distance	41
2.3.1	Mesure de recouvrement	41
2.3.2	Le rapport de distances	42
2.3.3	Distances Maximin et Minimax	42
2.3.4	Exemple	43
2.3.5	Algorithmes d'amélioration du critère de distance	44
2.4	Critère d'entropie	46
2.4.1	Définition de l'entropie	46
2.4.2	Entropie et quantité d'information.	47
2.4.3	Plans à entropie maximale	47
III PLANS D'EXPÉRIENCES PAR MÉTHODES MCMC		49
3.1	Introduction	49
3.2	Méthodes Monte Carlo	49
3.3	Méthodes McMC	51
3.4	Chaînes de Markov	52

3.4.1	Un cas simple dans un espace d'états discret	53
3.4.2	Réversibilité et mesure invariante	56
3.5	Processus ponctuels et planification d'expériences	57
3.5.1	Le processus ponctuel de Strauss	58
3.5.2	Extensions : processus ponctuel de Gibbs	60
IV	IMPLÉMENTATION, DISCUSIONS ET COMPARAISON . . .	63
4.1	Etude des plans de Strauss:	63
4.2	Etude des plans de Strauss-Gibbs:	69
4.3	Etude des plans minimax par Le rapport de distances:	73
4.4	Etude des plans à entropie maximale	75
	Conclusion	76
	Références	77

Remerciements

Nous remercions en premier lieu Dieu, le tout puissant de nous avoir donné tant de courage, de patience et de volonté pour l'élaboration de ce modeste travail.

Nos vifs et sincères remerciements à notre promoteur Mr : HICHEM EL MOSSAOUI qui a été à l'origine de ce projet, pour sa rigueur, son suivi, et ses précieux conseils.

Nous remercions vivement Mr : FAROUK HANNANE pour son encadrement et guidé tout le long de ce travail et pour les nombreux conseils, les informations et la bibliographie qu'il m'a fournis sur les plans d'expériences.

Nous tenons aussi à remercier monsieur le chef de département : TAMI OMAR ainsi que nos respectables professeurs et le personnel de la direction tous les membres du jury pour le temps qu'ils nous ont consacré. Enfin, un grand merci va également à tous les enseignants de département DE MATHEMATIQUES.

Principaux notions et symboles

k	Nombre des facteurs.
N	Nombre d'essais d'un plan d'expériences.
n	Nombre de points souhaités.
p	Nombre de coefficients d'un modèle mathématique.
ξ	Matrice d'expériences.
ξ_N	Matrice d'expériences à N expériences.
ξ^*	La matrice optimale.
A	variables d'origine
x	variable codée
y	Vecteur des réponses mesurées.
\hat{y}	Vecteur des réponses calculées.
a_0	Constant d'un modèle.
a_i	Coefficients des termes de premier degré.
a_{ii}	Coefficients des termes carrés.
a_{ij}	Coefficients des termes rectangles.
\hat{a}_i	Coefficients de la variable x_i dans un modèle déterminé par la méthode des moindres carrés
a	Vecteur de tous les coefficients réels du modèle.
\hat{a}	Vecteur de tous les coefficients déterminés avec l'hypothèse des moindres carrés.
e	Vecteur des résidus.
X	matrice du modèle
$({}^tXX)$	Matrice d'information.
$({}^tXX)^{-1}$	Matrice de dispersion.
$E(\hat{a})$	Espérance des coefficients
$var(\hat{a})$	Variance des coefficients

$d^2(y_u)$	Fonction de variance au point u .
$d(y_u)$	Fonction d'erreur de prédiction au point u .
$E f f - D$	Efficacité vis-à-vis du critère D
$.E f f - G$	Efficacité vis-à-vis du critère G
$E f f - M$	Efficacité vis-à-vis du critère M
$E f f - R$	Efficacité vis-à-vis du critère R
$D_n(X)$	discrépance extrême
$D_n^*(X)$	discrépance à l'origine
$DL2_n(X)^2$	discrépance extrême en norme 2
$DL2_n^*(X)^2$	discrépance à l'origine en norme 2
$dist(x^i, x^k)$	distance euclidienne entre deux points x^i et x^k
λ	Mesure de recouvrement
R	Le rapport de distances
D	l'ensemble de points du domaine expérimental
$H(X)$	L'entropie d'une variable aléatoire continue X à densité de probabilité f
$C(X)$	matrice de corrélation
$McMC$	Monte- Carlo par chaînes de Markov
$a(x, y)$	probabilité d'acceptation
γ	coefficient d'interaction entre les 2 expériences.
k	constante de normalisation
$s(x)$	potentiel global d'énergie
R	rayon
U	potentiels locaux
φ	fonction de potentiel
α	type de puissance

Liste des illustrations, graphiques et tableaux

Figure 1.1	Représentation du système
Figure 1.2	Le Domaine de variation du facteur
Figure 1.3	Représentation du domaine d'étude pour deux facteurs
Figure 1.4	L'ensemble des réponses qui correspond à tous les points du domaine d'étude et qui forme une surface de réponse
Tableau 1.1	matrices d'expériences
Tableau 1.2	plan d'expériences
Figure 2.1	Ellipsoïde de confiance
Figure 2.2	Les 80 premiers points en dimension 2 avec un sous-ensemble J défini par x et x' pour le calcul de la discrédance extrême
Figure 2.3	Un sous-ensemble J pour le calcul de la discrédance à l'origine
Figure 2.4	De la gauche vers la droite. Un plan aléatoire, et un plan factoriel complet à 49 points en dimension 2
Tableau 2.1	Discrédances en norme infinie pour $n=16, 25, 36$ et 49 pour les 2 types de plans suivants : aléatoires, plan factoriel en dimension 2
Tableau 2.2	Valeurs des critères de distance pour les 2 types de plans suivants : aléatoires, plan factoriel où $n=16, 25, 36$ et 49
Figure 3.1	Représentation des éléments de E dans $[0, 1]^2$
Figure 3.2	La fonction de potentiel puissance pour différents α et $R = 0.2$
Figure 4.1	un plan initial (aléatoire) et plan de Strauss obtenu à partir du plan initial avec la représentation des sphères de rayon $R/2$
Figure 4.2	A gauche, un plan de Strauss à 20 points en dimension 2 de paramètres $R = 0.1$ et $\gamma = 0.1$. A droite, un plan de Strauss de paramètres $R = 0.3$ et $\gamma = 0.1$.
Figure 4.3	Un plan de Strauss à 50 points avec $R = 0.8$ et $\gamma = 0.1$

- Figure 4.4** A gauche, un plan de Strauss de paramètres $R = 0.19$ et $\gamma = 0.5$
A droite, un plan de Strauss de paramètres $R = 0.19$ et $\gamma = 0.01$
- Figure 4.5** A gauche. Un plan initiale. A droite. Un plan de Strauss obtenu à partir d'un potentiel tout ou rien
- Figure 4.6** A gauche. Un plan initiale. A droite. Un plan de Strauss obtenu à partir d'un potentiel puissance $\alpha = 0.5$
- Figure 4.7** A gauche, un plan initiale pour $R = 19.72$ et A droite, un plan finale pour $R = 1.86$
- Figure 4.8** A gauche, un plan initiale pour $R = 18.2$ et A droite, un plan finale pour $R = 1.06$
- Figure 4.9** A gauche, un plan initial, aléatoire, à 20 points. A droite, un plan à 20 points construit en maximisant l'entropie

Liste des algorithmes

Algorithme 1. Algorithme pour générer un plan minimax par

Le rapport de distances

Algorithme 2. Algorithme pour générer un plan à entropie maximale

Algorithme 3. Algorithme de Métropolis-Hastings

**Algorithme 4. Algorithme de Metropolis-Hastings pour générer
des plans d'expériences dits de Strauss**

**Algorithme 5. Algorithme de Metropolis-Hastings pour générer des
plans d'expériences dits de Strauss_Gibb**

Introduction

Pour proposer une solution répondant aux objectifs industriels, il est parfois nécessaire de chercher l'information manquante en réalisant un ensemble d'expériences. Les décisions importantes prises à partir des résultats expérimentaux et le coût non négligeable d'une expérimentation interdisent de laisser à la seule intuition

de l'expérimentateur la recherche de la solution du problème.

Le présent travail s'inscrit dans le contexte de la simulation de phénomènes complexes par des codes de calculs très onéreux en temps de calcul. La simulation peut par ailleurs résoudre des problèmes simples d'une façon plus satisfaisante

qu'en utilisant les méthodes analytiques classiques. Il est par exemple tout à fait possible de construire en quelques jours un modèle de réservoir très simplifié pour évaluer une nouvelle découverte. L'utilisation des simulations se justifie par des choix économiques, la crédibilité des résultats, l'aide à la décision et la surveillance du champ. De nombreux phénomènes complexes de l'industrie sont modélisés par des codes de simulation numérique sophistiqués comme outil de prévision, d'optimisation et de décision. C'est le cas de la simulation numérique de production d'un réservoir pétrolier mais aussi de la simulation de crues et de calcul de rejet de CO₂, de la simulation d'une soufflerie à Mach 3, de la simulation d'accident nucléaire, de la simulation de crash automobile, etc. On s'intéresse ici tout particulièrement aux plans d'expériences numériques sur lesquels on évalue le simulateur et qui servent de base pour construire une approximation du code de calcul.

La différence fondamentale entre les plans d'expériences numériques et les plans classiques développés en laboratoire (expériences physiques) réside dans le fait qu'en laboratoire, on suppose la présence d'erreurs aléatoires dues aux mesures, aux conditions expérimentales, ... etc. Ce qui n'est en principe pas le cas avec

une expérience numérique (simulation déterministe). Nous n'envisageons pas ici le cas des simulateurs bruités, à savoir des simulations entachées de bruit de simulation

(Monte-Carlo ou autre). La présence d'une racine aléatoire dans le simulateur peut aussi induire des différences au niveau des réponses. Nous n'envisageons pas ce cas dans ce travail. Dans le cas général, les techniques des plans classiques telles que la randomisation et la répétabilité ne sont plus forcément de mise. En effet, les plans classiques ont tendance à positionner les points aux bords du domaine pour prendre en compte la variation aléatoire et offrir une tendance plus fiable en présence d'erreurs de mesures, on pense par exemple aux plans Box Behnken, plans composites, ... etc

Il existe cependant des attributs communs entre ces deux types de plans. Dans les deux cas, l'objectif principal est d'extraire le plus d'informations possibles avec un nombre d'expériences limité. Une autre caractéristique est que l'échantillonnage des points se fait indépendamment et donc se prête à une évaluation désordonnée. En numérique, il sera donc possible de faire des calculs parallèles ou bien d'utiliser des multiprocesseurs tout comme en expérimentation classique où il est possible de faire plusieurs manipulations en parallèle.

Le présent document se décompose en quatre chapitres.

Le premier chapitre fait un état de l'art sur les grands principes de la méthodologie de la recherche expérimentale (MRE).

Le deuxième chapitre étudie le bon remplissage de l'espace par des critères intrinsèques, nous discutons dans un premier temps les critères de discrédance pour mesurer si la distribution donnée est proche d'une distribution uniforme. Ensuite des critères de distance qui comparent une distribution donnée à une grille régulière. Enfin, nous présentons le critère d'entropie qui mesure la quantité d'information contenue dans le plan.

Le troisième chapitre présente une nouvelle méthode de planification d'expériences basée sur la méthodes de Monte-Carlo par chaînes de Markov. Nous présentons également dans cette partie des extensions qui permettent de traiter des problèmes particuliers tels que la génération de simulation dans un domaine contraint.

Enfin dans le dernier chapitre nous présentons les différents résultats obtenus par l'implémentation des différentes méthodes traitées dans les deuxième et troisième chapitres avec une étude de comparaison entre le plan de Strauss et celui de Gibbs.

CHAPITRE I

METHODOLOGIE DE LA RECHERCHE EXPERIMENTALE

Les plans d'expériences sont des manipulations réalisées à des fins scientifiques dans des conditions rigoureusement contrôlées. Ce chapitre traite les plans d'expériences d'une façon générale dont il parle de leurs intérêts, leurs différents types ainsi que les étapes à suivre pour réaliser une expérience et à la fin il parle des choix d'expériences.

1.1 Historique

La méthode des **plans d'expériences** est à la fois nouvelle et ancienne . Les scientifiques n'ont abordé ce sujet que depuis peu d'années. Les premiers qui se sont penchés sur ce problème sont des agronomes et des statisticiens. Les techniques et les notions qu'ils ont développées sont si générales qu'elles peuvent être utilisées dans beaucoup de domaines. Cette méthode est récente puisqu'on peut la faire démarrer avec les travaux de **R.A .Fisher** (début de vingtième siècle). Aussi bizarre que cela paraisse, elle ne porte pas encore de nom. Les chercheurs ont proposés **Expérimentique** ou **Expérimentologie**. Après 1945, les plans d'expériences suscitent de nombreuses publications et recherches dans le monde anglo-saxon. Des statisticiens comme **Yate, Youden, Cochran, Plackett et Burman**, et bien d'autres, enrichissent et divulguent la méthode. **Box et Hunter**, s'appuyant notamment sur les travaux de Yate, développent des méthodes particulières de construction de **plans fractionnaires à deux niveaux**. A partir des années 50, en s'attaquant à l'amélioration de la qualité, le Japon imprime un nouveaux souffle. **Taguchi et Masuyama** élaborent des tables permettant de construire des **plans d'expériences**

orthogonaux adaptés à la majorité des problèmes industriels. Ces tables ont été publiées en 1959 et 1961. Par la suite les méthodes initiées par Taguchi se sont répandues aux Etats-unis puis en Europe. Il faut signaler l'existence sur le marché de logiciels de plus en plus conviviaux destinés à la construction et à l'analyse des plans d'expériences. Il faut noter cependant que si ces logiciels déchargent l'utilisateur de certains calculs parfois fastidieux, ils ne permettent pas de poser le problème. En d'autres termes les logiciels ne doivent être utilisés qu'après une solide formation à la méthode. Autrement les risques de fausses conclusions seront importants.[5]

1.2 Intérêts des plans d'expériences

Les plans d'expériences sont utilisés dans les études industrielles en recherche-développement.

Ils interviennent dans de nombreux domaines industriels, comme:[1]

- Industries chimiques, pétrochimiques et pharmaceutiques
- Industries mécaniques et automobiles
- Industries métallurgiques . . . etc

Leur utilisation vise aux buts suivants :

- Détermination des facteurs clés dans la conception d'un nouveau produit ou d'un nouveau procédé
- Optimisation des réglages d'un procédé de fabrication ou d'un d'appareil de mesure
- Prédiction par modélisation du comportement d'un procédé

Les plans d'expériences s'inscrivent dans une démarche générale d'amélioration de la qualité.

Le succès de la démarche originale des plans d'expériences réside dans la possibilité d'interprétation de résultats expérimentaux avec un effort minimal sur le plan expérimental, ce pendant la minimisation du nombre nécessaire d'expériences permet un gain en temps et en coût financier.

1.3 Notion d'espace expérimental

Le scientifique est souvent amené à comprendre comment réagit un système en fonction des facteurs susceptibles de le modifier. Pour visualiser cette évolution, il mesure une réponse et va ensuite essayer d'établir des relations de cause à effet entre les réponses et les facteurs.[8]

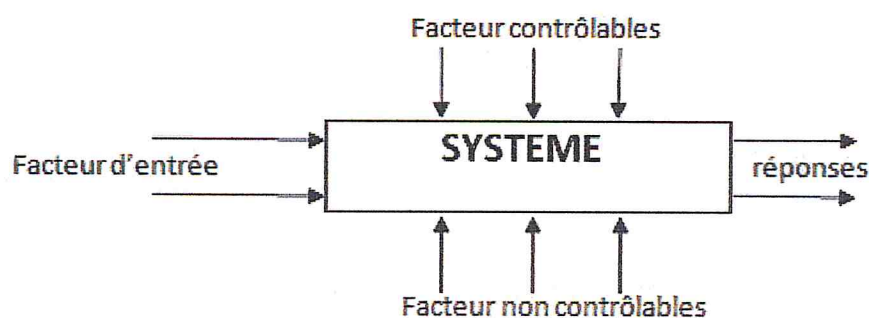


Figure 1.1:Représentation du système

Parmi les facteurs on distinguera:

- les facteurs contrôlables qui dépendent directement du choix du technicien (pression, température, matériau ... etc)
- les facteurs non contrôlables qui varient indépendamment du choix du technicien (conditions climatiques, environnement d'utilisation ... etc)

La grandeur d'intérêt mesurée à chaque essai est appelée **la réponse**. La valeur de cette réponse, dépend de plusieurs variables. Au lieu de terme variable nous emploierons le mot **facteur** .

Le premier facteur peut être représenté par un axe gradué et orienté (Figure 1.2). la valeur donnée à un facteur pour réaliser un essai est appelée **niveaux**. Lorsque on étudie l'influence d'un facteur, en général, on limite ses variations entre deux bornes (borne inférieure c'est le **niveau bas** et borne supérieure c'est le **niveau haut**).

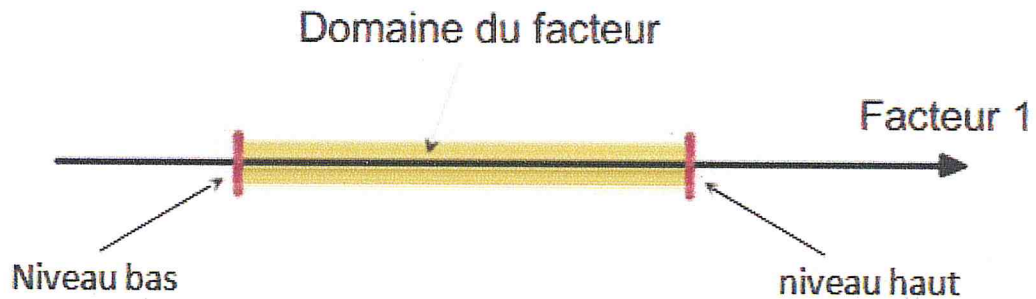


Figure 1.2 : Le Domaine de variation du facteur

1.3.1 Domaine d'étude (DE)

Etant donnée la définition des variables de type facteur, il devient naturel de définir un espace k dimensionnel, dans lequel chaque point représente une combinaison des valeurs possibles des k facteurs.

Cet espace est appelé **Domaine d'étude**. On trouve l'appellation **espace de recherche**, montrant un lien plus fort avec les futures optimisations prenant place dans cet espace - [6]

Les points intérieurs à ce domaine sont les **points d'expérience**, car ils sont la traduction mathématique, par leurs coordonnées, des combinaisons de valeurs des grandeurs physiques leur correspondant.

Le **Domaine d'étude** est obligatoirement de forme hyper-rectangulaire, ses dimensions étant imposées par les plages de variation des k facteurs.

Ainsi, pour $k = 2$, on peut représenter le **Domaine d'étude** comme une zone rectangulaire limitée par les bornes prises par les deux facteurs (x_1 et $x_2 \in [-1, +1]$)

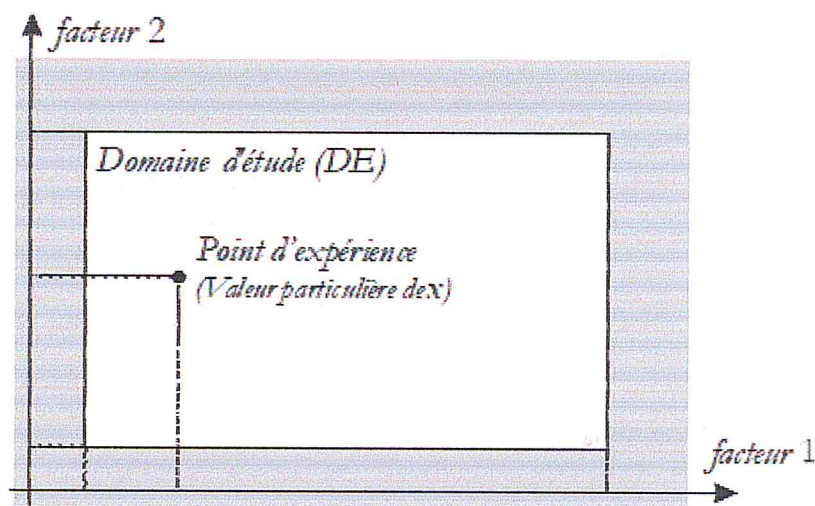


Figure 1.3: Représentation du domaine d'étude pour deux facteurs

Ainsi tout point de cet espace k -dimensionnel est défini par ses coordonnées, c'est-à-dire par un vecteur x à k éléments, x servira de variable de type inconnu pour l'écriture des modèles exprimant les variations des réponses en fonction des valeurs des facteurs, un modèle est donc une relation simple de la forme :

$$y = f(x)$$

Où y est une réponse en particulier, x prenant ses valeurs dans le Domaine d'étude.

1.3.2 Surfaces de réponse

À chaque point du domaine d'étude correspond une réponse. À l'ensemble de tous les points du domaine d'étude correspond un ensemble de réponses qui se localise sur une surface appelée surface de réponse (Figure 1.4). [6]

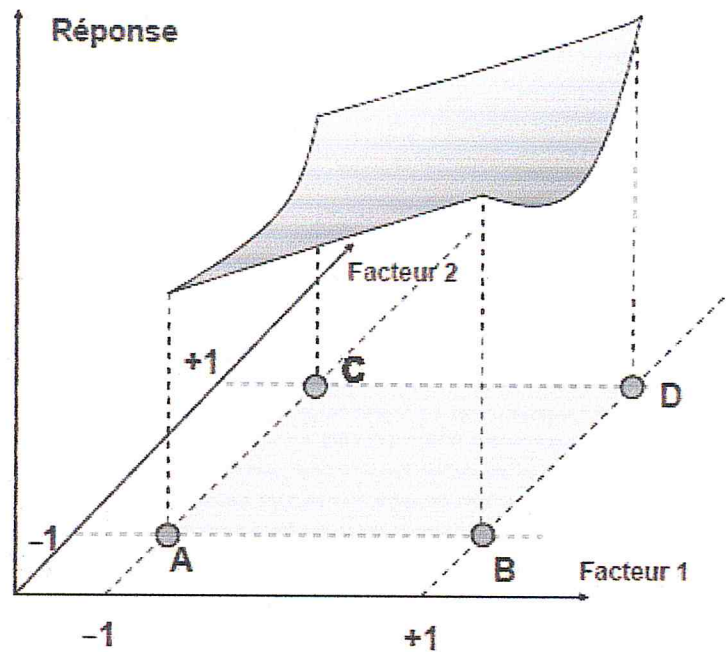


Figure 1.4 - L'ensemble des réponses qui correspond tous les points du domaine d'étude et qui forme une surface de réponse.

En général, on ne connaît que quelques réponses, celles qui correspondent aux points expérimentaux retenus par l'expérimentateur. On interpole à l'aide d'un modèle mathématique, les réponses inconnues pour obtenir la surface de réponse.

Les points d'expériences retenus par la théorie des plans d'expériences assurent la meilleure précision possible sur la forme et la position de la surface de réponse.

1.3.3 Variables centrées réduites

Lorsqu'on attribue la valeur -1 au niveau bas d'un facteur et la valeur $+1$ au niveau haut, on effectue deux modifications importantes :

- On déplace l'origine des mesures. Dans l'exemple choisi, le milieu de l'intervalle $[-1, +1]$ correspond à une valeur de 100 km/h. La valeur numérique de la nouvelle origine, zéro, diffère donc de l'origine exprimée en unité courante.

• On change l'unité de mesure. Par exemple, si le niveau bas du facteur « vitesse du véhicule » est 80 km/h et le niveau haut 120 km/h, il y a 40 km/h entre ces deux valeurs, soit 40 fois l'unité de vitesse. Entre -1 et $+1$ il y a deux unités nouvelles : la nouvelle unité vaut 20 km/h, on lui donne le nom de **Pas**.

Ces deux modifications entraînent l'introduction de nouvelles variables que l'on appelle variables centrées réduites (v.c.r) , centrées pour indiquer le changement d'origine et réduites pour signaler la nouvelle unité. On utilise également le terme de variables codées ou d'unités codées.

Le passage des variables d'origine A aux variables codées x , et inversement, est donné par la formule suivante (A_0 est la valeur centrale en unités courantes) :

$$x = \frac{A - A_0}{pas} \quad pas = \frac{A_{+1} - A_{-1}}{2} \quad A_0 = \frac{A_{+1} + A_{-1}}{2} \quad (2)$$

1.3.4 Matrice d'expérience

Nous allons illustrer ci-dessous la notion de matrice d'expériences avec un exemple simple. Prenons deux facteurs quantitatifs A_1 et A_2 . Le facteur A_1 peut varier par exemple entre 0.5 et 2 , le facteur A_2 peut varier entre 100 et 200. Supposons que la matrice d'expériences choisie soit celle reportée dans le tableau 1-1. [7]

Numéro d'essai	A_1	A_2
1	0.5	100
2	0.5	200
3	2	100
4	2	200

Tableau 1-1 : matrice
d'expériences

Ainsi on code le tableau 1-1 (matrice d'expériences) à l'aide de la relation (2) pour obtenir le plan d'expériences à réaliser (tableau 1-2).

Numéro d'essai	X_1	X_2
1	-1	-1
2	-1	1
3	1	-1
4	1	1

Tableau 1-2 : plan d'expériences

1.4 Modélisation mathématique a priori de la réponse [5]

1.4.1 Modélisation mathématique

En l'absence de toute information sur la fonction qui lie la réponse aux facteurs, on se donne a priori une loi d'évolution dont la formulation la plus générale est la suivante :

$$y = f(x_1, x_2, x_3, \dots, x_n) \quad (1)$$

Cette fonction est trop générale et il est d'usage d'en prendre un développement limité de Taylor-Mac Laurin, c'est-à-dire une approximation. Si les dérivées peuvent être considérées comme des constantes, le développement précédent prend la forme d'un polynôme de degré plus ou moins élevé :

$$y = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j + \sum a_{ii} x_i^2 + \dots \quad (2)$$

où :

- y est la grandeur à la quelle s'intéresse l'expérimentateur ; c'est la réponse ou la grandeur d'intérêt.

- x_i représente un niveau du facteur i .
- x_j représente un niveau du facteur j .
- a_0, a_i, a_{ij}, a_{ii} sont les coefficients du polynôme.

Ce modèle est appelé le modèle a priori ou le modèle postulé.

Les modèles établis sont des modèles de prévision valables dans le domaine d'étude, domaine que l'on doit toujours préciser. Ces modèles ne sont pas des modèles théoriques basés sur des lois physico-chimiques ou mécaniques. Dans quelques rares cas, il est possible d'utiliser des lois physiques théoriques connues.

1.4.2 Modélisation expérimentale

Deux compléments doivent être apportés au modèle purement mathématique précédemment décrit (l'équation (2)), Le premier complément est le manque d'ajustement. Cette expression traduit le fait que le modèle choisi par l'expérimentateur avant les expériences est probablement un peu différent du modèle réel qui régit le phénomène étudié. Il y a un écart entre ces deux modèles. Cet écart est le manque d'ajustement, on le note par la lettre Δ .

Le second complément est la prise en compte de la nature aléatoire de la réponse.

En effet, dans le cas général, si l'on mesure plusieurs fois une réponse en un même point expérimental, on n'obtiendra pas exactement le même résultat. Il y a une

dispersion des résultats. Les dispersions ainsi constatées sont appelées erreurs aléatoires ou erreurs expérimentales et on les note par la lettre σ_y .

La relation générale (1) doit être modifiée ainsi :

$$y = f(x_1, x_2, x_3, \dots, x_n) + \Delta + \sigma_y$$

1.4.3 Modèle du statisticien

Le modèle précédent est trop compliqué pour faire des calculs. On effectue les simplifications suivantes :

- L'erreur d'ajustement entre le modèle choisie à priori et le modèle réel est supposée négligeable
- Les erreurs aléatoires sont issues d'une même population, pour tout le domaine d'étude. Il s'agit de valeurs ayant la même moyenne, le même écart type et provenant de la même distribution. Il faut bien voir que les valeurs de l'erreur sont différentes les unes des autres, puisqu'elles résultent de tirages au hasard dans une population donnée. Pour un modèle polynomial du second degré à deux facteurs, le modèle de statisticien s'écrit :

$$y = f(x_1, x_2, x_3, \dots, x_n) + \sigma_y$$

1.4.4 Système d'équations

Chaque point expérimental apporte une valeur de la réponse. Or cette réponse est modélisée par un polynôme dont les coefficients sont les inconnues qu'il faut déterminer. À la fin du plan d'expériences, on a un système de n équations (s'il y a n essais) à p inconnues (s'il y a p coefficients dans le modèle choisi *a priori*). Ce système s'écrit d'une manière simple en notation matricielle :

$$y = Xa + \varepsilon$$

où :

- y est le *vecteur des réponses*,
- X est la *matrice de calcul des coefficients* ou *matrice du modèle* qui dépend des points expérimentaux choisis pour exécuter le plan et du modèle postulé,
- a est le *vecteur des coefficients*,
- ε est le *vecteur des écarts*

1.4.5 Estimation des coefficients par la méthode des moindres carrés

Lorsque on estime les inconnues a_0, a_1, \dots, a_q par $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_q$ nous pouvons calculer la réponse au point i par :

$$\hat{y}_i = \hat{a}_0 + \hat{a}_1 x_{i1} + \hat{a}_2 x_{i2} + \dots + \hat{a}_j x_{ij} + \dots + \hat{a}_q x_{iq}$$

La valeur \hat{y}_i diffère du résultat expérimental y_i de la quantité e_i (de même que les \hat{a}_i sont les estimateurs de a_i , les e_i , sont des estimations des ε_i).

$$y_i = \hat{y}_i + e_i$$

Nous écrirons cette égalité quelque soit i , nous obtenons le système linéaire :

$$y_1 = \hat{a}_0 + \hat{a}_1 x_{11} + \hat{a}_2 x_{12} + \dots + \hat{a}_j x_{1j} + \dots + \hat{a}_q x_{1q} + e_1$$

$$y_2 = \hat{a}_0 + \hat{a}_1 x_{21} + \hat{a}_2 x_{22} + \dots + \hat{a}_j x_{2j} + \dots + \hat{a}_q x_{2q} + e_2$$

$$y_i = \hat{a}_0 + \hat{a}_1 x_{i1} + \hat{a}_2 x_{i2} + \dots + \hat{a}_j x_{ij} + \dots + \hat{a}_q x_{iq} + e_i$$

$$y_N = \hat{a}_0 + \hat{a}_1 x_{N1} + \hat{a}_2 x_{N2} + \dots + \hat{a}_j x_{Nj} + \dots + \hat{a}_q x_{Nq} + e_N$$

Nous cherchons les valeurs des \hat{a}_j qui minimisent la somme des carrés des écarts $\sum e_i^2$. Concrètement nous cherchons le modèle linéaire qui passe au plus près de l'ensemble des points expérimentaux.

Pour simplifier l'écriture, nous adoptons la notation matricielle :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1j} & \dots & x_{1q} \\ 1 & x_{21} & \dots & x_{2j} & \dots & x_{2q} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ 1 & x_{i1} & \dots & x_{ij} & \dots & x_{iq} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ 1 & x_{N1} & \dots & x_{Nj} & \dots & x_{Nq} \end{pmatrix} \begin{pmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \vdots \\ \hat{a}_j \\ \vdots \\ \hat{a}_q \end{pmatrix} + \begin{pmatrix} e_0 \\ e_1 \\ \vdots \\ e_i \\ \vdots \\ e_N \end{pmatrix}$$

Le système à résoudre s'écrit : $y = X\hat{a} + e$, et le critère des moindres carrés ${}^t ee$ est minimale si on a :

$$\begin{aligned} {}^t ee &= {}^t (y - X\hat{a})(y - X\hat{a}) \\ &= ({}^t y - {}^t \hat{a} {}^t X)(y - X\hat{a}) \\ &= {}^t yy - {}^t \hat{a} {}^t Xy - {}^t yX\hat{a} + {}^t \hat{a} {}^t XX\hat{a} \\ &= {}^t yy - 2{}^t \hat{a} {}^t Xy + {}^t \hat{a} {}^t XX\hat{a} \end{aligned}$$

Calculons la dérivée de ${}^t ee$ par rapport à l'inconnue

$$\frac{\partial ({}^t ee)}{\partial \hat{a}} = \frac{\partial ({}^t yy)}{\partial \hat{a}} - 2 \frac{\partial ({}^t \hat{a} {}^t Xy)}{\partial \hat{a}} + \frac{\partial ({}^t \hat{a} {}^t XX\hat{a})}{\partial \hat{a}}$$

Où:

$$\frac{\partial ({}^t yy)}{\partial \hat{a}} = 0$$

car ${}^t yy$ ne dépend pas de \hat{a}

$$\frac{\partial ({}^t \hat{a} {}^t Xy)}{\partial \hat{a}} = {}^t Xy$$

car ${}^t \hat{a} {}^t Xy$ est une forme linéaire en \hat{a}

$$\frac{\partial ({}^t \hat{a} {}^t XX\hat{a})}{\partial \hat{a}} = 2{}^t XX\hat{a}$$

car ${}^t \hat{a} {}^t XX\hat{a}$ est une forme quadratique en \hat{a}

Il vient donc :

$$\frac{\partial ({}^t ee)}{\partial \hat{a}} = -2 {}^t Xy + 2 {}^t XX\hat{a}$$

La valeur de \hat{a} qui minimise ${}^t ee$ doit vérifier :

$$\begin{aligned} \frac{\partial ({}^t ee)}{\partial \hat{a}} &= 0 \implies -2 {}^t Xy + 2 {}^t XX\hat{a} = 0 \\ &\implies {}^t XX\hat{a} = {}^t Xy \end{aligned}$$

Si la matrice $({}^tXX)^{-1}$ n'est pas singulière on a :

$$\hat{a} = ({}^tXX)^{-1} {}^tXy$$

Cette relation est fondamentale et nous l'utiliserons constamment par la suite. Elle est valable pour tous les modèles polynomiaux quelque soit leur degré et quel que soit le nombre de coefficients. Les coefficients ainsi obtenu sont utilisés pour écrire le modèle mathématique. Ce modèle permet de calculer les réponses (qui sont souvent appelées les réponses prédites \hat{y}) dans tout le domaine d'étude. [10]

$$\hat{y} = X\hat{a}$$

1.4.6 Propriétés des coefficients

les coefficients \hat{a} sont des résultats expérimentaux et doivent être considérés comme des variables aléatoires, donc nous pouvons calculer leur espérance mathématique ainsi que leur variance.

1.4.6.1 Espérance mathématique des coefficients

l'espérance mathématique de \hat{a} a pour expression :

$$\begin{aligned} E(\hat{a}) &= E\left[({}^tXX)^{-1} {}^tXy\right] \\ &= ({}^tXX)^{-1} {}^tXE(y) \end{aligned}$$

car les éléments de X sont considérer comme constantes. En désignant par a le vecteur des coefficients vrais et ε le vecteur des N écarts entre les résultats expérimentaux et les réponses théoriques alors :

$$\begin{aligned} y &= X.a + \varepsilon \\ E(y) &= E(X.a + \varepsilon) = E(X.a) \\ &= X.E(a) \end{aligned}$$

car $E(\varepsilon) = 0$. nous trouvons:

$$E(\hat{a}) = ({}^tXX)^{-1} {}^tXXa = a$$

le résultat que nous venons d'établir signifie que les distributions des \hat{a}_i sont centrés sur les valeurs vraies a_i (a_i sans biais)

1.4.6.2 Variance des coefficients

par définition la variance de \hat{a} est

$$\text{var}(\hat{a}) = E[(\hat{a} - a)^t(\hat{a} - a)]$$

Remplaçons \hat{a} par $({}^tXX)^{-1} {}^tXy$ et y par $X.a + \varepsilon$. Nous obtenons :

$$(\hat{a} - a) = ({}^tXX)^{-1} {}^tX(X.a + \varepsilon) - a = a + ({}^tXX)^{-1} {}^tX\varepsilon - a = ({}^tXX)^{-1} {}^tX\varepsilon$$

d'où

$${}^t(\hat{a} - a) = \varepsilon^t X ({}^tXX)^{-1}$$

donc,

$$\text{var}(\hat{a}) = E\left[({}^tXX)^{-1} X^t \varepsilon^t \varepsilon X ({}^tXX)^{-1}\right] = ({}^tXX)^{-1} {}^tX E(\varepsilon^t \varepsilon) X ({}^tXX)^{-1}$$

Remplaçons $E(\varepsilon^t \varepsilon)$ par $E[(\varepsilon - 0)^t(\varepsilon - 0)] = \text{var}(\varepsilon) = \sigma^2$. Nous pouvons écrire:

$$\begin{aligned} \text{var}(\hat{a}) &= ({}^tXX)^{-1} {}^tX \sigma^2 X ({}^tXX)^{-1} \\ &= \sigma^2 ({}^tXX)^{-1} {}^tXX ({}^tXX)^{-1} \\ \implies \text{var}(\hat{a}) &= \sigma^2 ({}^tXX)^{-1} \end{aligned}$$

En multipliant les éléments de la matrice $({}^tXX)^{-1}$ par σ^2 , Nous obtenons la matrice de variance-covariance suivante:

$$\text{var}(\hat{a}) = \begin{pmatrix} \text{var}(\hat{a}_0) & \dots & \dots & \text{cov}(\hat{a}_q, \hat{a}_0) \\ & \text{var}(\hat{a}_i) & \text{cov}(\hat{a}_j, \hat{a}_i) & \\ & \text{cov}(\hat{a}_i, \hat{a}_j) & \text{var}(\hat{a}_j) & \\ \text{cov}(\hat{a}_0, \hat{a}_q) & & & \text{var}(\hat{a}_q) \end{pmatrix}$$

la matrice de variance -covariance ci-dessus est comme $({}^tXX)^{-1}$, une matrice carrée symétrique. les éléments de la diagonale sont les variances des coefficients, tandis que les termes hors diagonale correspondent aux covariances.[9]

1.5 Plans classiques

Nous avons choisi de discuter des plans classiques dans ce mémoire car ils ont été développés à l'origine pour des applications des surfaces de réponses. Parmi ces plans, les plus courants sont les plans factoriels, les plans de Box-Behnken, composites centrés, plan Dohrlert,...etc

1.5.1 Plans de criblages

Ces plans permettent de découvrir les facteurs les plus influents sur une réponse donnée. On ne cherche pas vraiment à obtenir une relation précise entre les variations des facteurs et celle de la réponse. Ce dernier a comme type :

- **les plans factoriels complets 2^k à deux niveaux:** ces plans sont gourmands en essais dès que l'on dépasse trois facteurs. ils peuvent servir à faire des criblages mais aussi à faire de la modélisation.

- **les Plans factoriels fractionnaires 2^{k-p} :** Ce type permet d'étudier beaucoup de facteurs, si on sait bien les interpréter. On peut détecter les interactions entre les facteurs.

- **les Plans sursaturés :** Sont des plans qui prennent en compte un grand nombre de facteurs et qui ne demandent que quelques essais.

1.5.2 Plans de modélisation

• **les Plans factoriels complets à trois niveaux** : Ces plans permettent d'étudier k facteurs à raison de trois niveaux par facteur. La notation de ces plans est 3^k , le k en exposant signifie qu'il y a k facteurs à étudier, et le trois signifie que chaque facteur prend 3 niveaux. Cette notation indique également le nombre d'essais à effectuer.

• **les Plans composites** : Un plan composite est constitué de trois parties :

★ Un plan factoriel dont les facteurs prennent deux niveaux.

★ Au moins un point expérimental situé au centre du domaine d'étude.

★ Des points **axiaux**. Ces points expérimentaux sont situés sur les axes de chacun des facteurs.

• **les Plans de Doehlert** : La caractéristique principale des plans de Doehlert est d'avoir une répartition uniforme des points expérimentaux dans l'espace expérimental. la disposition de ces points pour un plan à deux facteurs (essais 1 à 7). Tous les points sont à la même distance du centre du domaine d'étude et sont situés sur le cercle trigonométrique. Ils forment un hexagone régulier.

• **Les plans de Box-Behnken** : Box et Behnken ont proposé en 1960 ces plans qui permettent d'établir des modèles du second degré. En effet tous les facteurs ont trois niveaux, -1 , 0 et $+1$. Ces plans sont faciles à mettre en oeuvre.

1.5.3 Plans de mélanges

Les plans de mélanges sont de plans que l'on utilise lorsqu'on étudie des produits composés de plusieurs constituants, l'objectif est de trouver la loi qui réagit une ou plusieurs réponses en fonction de la composition du mélange.

CHAPITRE II

CRITÈRES ET PLANS OPTIMAUX

2.0.3.1 Introduction

Etudier l'uniformité d'une distribution de points est un travail fastidieux, voire même impossible en grande dimension. Il est donc utile de se référer à des critères afin de pouvoir décider si une distribution est uniforme ou non et si elle vérifie également le bon remplissage de l'espace. Il existe aujourd'hui de nombreux critères d'optimalités.

Nous discutons dans un premier temps les critères de discrédance qui permettent de mesurer si la distribution donnée est proche d'une distribution uniforme. Ensuite, nous introduisons la classe des critères de distance qui comparent une distribution donnée à une grille régulière.

Enfin, nous présentons le critère d'entropie qui mesure la quantité d'information contenue dans le plan.

Dans chacune des parties, nous discutons de plans optimaux au sens de ces critères, choisis pour leur bonne qualité d'uniformité ou de remplissage. Aussi, on y retrouve les plans à faible discrédance, des plans améliorés à l'aide de critères de distance, et enfin des plans où l'on maximise l'entropie associée.

2.1 Critères d'optimalités classique

Il existe aujourd'hui de nombreux critères d'optimalité. Il serait difficile d'être exhaustif, mais on essaiera toute fois d'en présenter une liste suffisamment conséquente pour bien apprécier la variété des choix possibles et les principaux critères d'optimalité classique sont

2.1.1 Qualité de la représentation des coefficients

Plusieurs auteurs ont montré que la limite du domaine de confiance des coefficients est donnée par la relation

$${}^t(a - \hat{a}) ({}^tXX) (a - \hat{a}) = qs^2 \quad (1)$$

Où

s^2 est une estimation de la variance expérimentale

F est la statistique correspondant au test de Fisher

α est le niveau de confiance choisi

ddl est le nombre de degrés de liberté associé à s^2

Ce domaine est un hyper ellipsoïde dans l'espace des coefficients, centré sur le vecteur calculé \hat{a} . Avec un risque α , on peut dire, que les valeurs vraies des coefficients sont dans cet hyper ellipsoïde. Pour deux coefficients, l'hyper ellipsoïde est représenté graphiquement par :

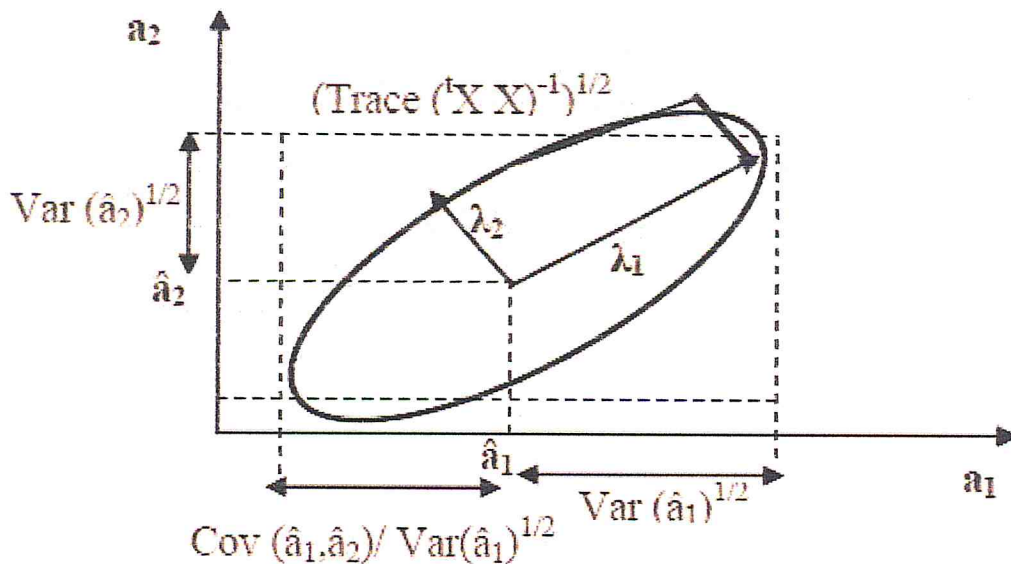


Figure 2.1 : Ellipsoïde de confiance

Où (\hat{a}_1, \hat{a}_2) sont les deux coefficients trouvés par la régression multilinéaire, λ_1 et λ_2 les deux valeurs propres de l'ellipsoïde. Cet hyper ellipsoïde représente les performances de l'ensemble des observations pour un modèle donné. Il peut être décrit par 3 caractérisations :

A. Volume

Le volume de l'ellipsoïde est égal à $\pi (\lambda_1, \lambda_2)^{1/2} = \pi \left(\det ({}^tXX)^{-1} \right)^{1/2}$. Ce volume est lié au déterminant de la matrice de dispersion $({}^tXX)^{-1}$. Plus le volume sera petit, plus l'hyper ellipsoïde tend vers un point et plus on se rapprochera de la solution vraie A.

B. Forme

Lorsque la forme est très allongée, alors il existe une grande disparité sur la précision des coefficients. Si $\lambda_1 = \lambda_2$ l'ellipsoïde devient une hyper sphère, et donc tous les coefficients sont déterminés avec la même précision.

C. Orientation

Si les axes principaux de l'hyper ellipsoïde sont parallèles aux axes des coefficients alors, la valeur calculée d'un coefficient sera indépendante des valeurs calculées des autres coefficients.

D'après la formule (1), l'hyper ellipsoïde dépend de la matrice de $({}^tXX)$, donc les critères sont fondés sur les matrices d'information $({}^tXX)$ et de dispersion $({}^tXX)^{-1}$. [5]

2.1.2 Critère A

Une matrice d'expériences est dite **A-optimale** si elle conduit à la trace minimale de $({}^tXX)^{-1}$ matrice de dispersion.

2.1.3 Critère D

Le critère-D est défini par :

$$D(\xi^*) = \left| ({}^tXX)^{-1} \right|$$

Une matrice d'expériences ξ^{*D} est dite **D-optimale** si elle conduit au déterminant minimal pour sa matrice de dispersion.

$$\xi^{*D} = \arg_{\xi \in E} \min D(\xi)$$

2.1.3.1 Efficacité-D

L'efficacité de la matrice d'expériences ξ^1 par rapport à la matrice d'expériences ξ^2 est définie par :

$$\begin{aligned} \text{Eff-D}(\xi^1, \xi^2) &= \left(\frac{|({}^t X_1 X_1)^{-1}|}{|({}^t X_2 X_2)^{-1}|} \right) \cdot 100 \\ &= \left(\frac{|({}^t X_2 X_2)|}{|({}^t X_1 X_1)|} \right) \cdot 100 \end{aligned}$$

2.1.4 Critère E

Une matrice d'expériences est dite **E-optimale** si la valeur propre maximale de $({}^t X X)^{-1}$ est la plus faible possible

2.1.5 Critère G

Ce critère prend en compte la plus grande valeur sur tout le domaine expérimental, de la fonction de variance $d^2(y_u)$ engendré par la matrice ξ_N .

$$d_{\max}(\xi_N) = \max(d^2(y_u))$$

La meilleure matrice d'expériences vis à vis de ce critère étant celle qui a la plus petite valeur

2.1.5.1 Efficacité-G

L'efficacité-G de la matrice d'expériences ξ^2 par rapport à la matrice d'expériences ξ^1 est définie par

$$\text{Eff-G}(\xi^1, \xi^2) = \left[\frac{d_{\max}(\xi^1)}{d_{\max}(\xi^2)} \right] \cdot 100$$

Avec: $d_{\max}(\xi^1) \leq d_{\max}(\xi^2)$.

2.1.6 Critère M

le critère M permet de tenir compte de la qualité d'information apportée par l'expérience.

Ce critère est indépendant du nombre des essais du plan d'expériences. On appelle matrice des moments M , la matrice définie par :

$$M(\xi_N) = \frac{{}^t X X}{N}$$

Pour une matrice d'expériences $\xi_N \in E$, le critère M est défini par :

$$M(\xi_N) = |M(\xi_N)| = \frac{|{}^t X X|}{N^p}$$

où p est l'ordre de la matrice $({}^t X X)$

Une matrice d'expériences ξ^{*M} est dite M optimale si :

$$\xi_N^{*R} = \arg_{\xi \in E} \max(M(\xi_N))$$

2.1.6.1 Efficacité-M

Soient M_1 et M_2 deux matrices des moments associés à deux matrices d'expériences 1 et 2 constituées respectivement de N_1 et N_2 expériences. L'efficacité 1 par rapport à 2 selon le critère M est définie par :

$$\begin{aligned} \text{Eff} - M(\xi^1 \xi^2) &= \left(\frac{|M(\xi^1)|}{|M(\xi^2)|} \right)^{1/p} \cdot 100 \\ &= \frac{N_2}{N_1} \left(\frac{|{}^t X_1 X_1|}{|{}^t X_2 X_2|} \right) \cdot 100 \end{aligned}$$

2.1.7 Critère-R

Pour une matrice d'expériences ξ_N le critère R est défini par

$$R(\xi_N) = N$$

La minimisation de ce critère n'est pas une fin en soi. Il s'agit plutôt de rechercher le nombre minimal d'expériences permettant d'atteindre les objectifs fixés par l'expérimentateur.

2.1.7.1 Efficacité-R

Soit $\xi_{N_1} \in E$ et $\xi_{N_2} \in E$, telles que $N_1 < N_2$. Alors, l'efficacité R de la matrice d'expériences ξ_{N_1} vis-à-vis de la matrice d'expériences ξ_{N_2} est définie par :

$$Eff - R(\xi_1, \xi_2) = \frac{N_1}{N_2} \cdot 100$$

Plus $Eff - R$ est faible, meilleure est la matrice d'expériences ξ_{N_1} par rapport à la matrice d'expériences ξ_{N_2} , suivant ce critère.

L'efficacité-R de ξ_N par rapport à la matrice d'expériences optimale-R ξ^{*R} est :

$$Eff - R(\xi^{*R}, \xi_N) = \frac{p}{N} \cdot 100$$

Avec p = nombre de coefficients du modèle, i.e. nombre minimal d'expériences.

D'où

$$\xi^{*R} = \xi_p$$

2.1.8 Critère d'orthogonalité

Une matrice d'expériences est orthogonale si elle permet d'obtenir des estimations des coefficients indépendants. Cela se caractérise par des axes de l'ellipsoïde parallèles aux axes des coefficients. Cette propriété est obtenue quand $({}^tXX)^{-1}$ est diagonale donc quand les covariances des coefficients sont nulles.

2.1.9 Critère de Presque Orthogonalité

Si la matrice $({}^tXX)^{-1}$ obtenue en retirant sa première ligne et sa première colonne est diagonale, le critère de presque orthogonalité est respecté.

2.1.10 Critère d'isovariation par rotation

On désire que les réponses calculées avec le modèle issu du plan d'expériences aient une erreur de prévision identique pour des points situés à la même distance du centre du domaine d'étude. Dans ce cas on parle de plan iso variant par rotation.

2.2 critère d'uniformité basé sur la discrédance

2.2.1 La discrédance

consiste à mesurer l'écart entre une distribution de points donnée et une distribution de points uniforme ; en d'autres termes, elle mesure l'irrégularité de la distribution. On donne ci-dessous la définition permettant de comprendre le principe sur lequel elle est basée.

Définition:

Soit X la suite de n points $x^1, x^2, x^3, \dots, x^n$ dans $[0, 1]^d$ et soit J un sous-ensemble de $[0, 1]^d$

avec les notations précédentes on pose:

$$D_n(J, X) = \frac{A(J, X)}{N} - \lambda_d(J)$$

où $A(J, X)$ est le nombre d'indices i , $1 \leq i \leq n$, tel que $x^i \in J$ et $\lambda_d(J)$ est la mesure de Lebesgue (ou volume) de J . on distingue 2 type de discrédance

- La **discrédance extrême** de X , notée $D_n(X)$ est définie par

$$D_n(X) = \sup_{J \in \mathfrak{S}} |D_n(J, X)|$$

où \mathfrak{S} est l'ensemble des parties de $[0, 1]^d$ de la forme $\prod_{i=1}^d [a_i, b_i]$.

- La **discrédance à l'origine** de X , notée $D_n^*(X)$ est définie par

$$D_n^*(X) = \sup_{J \in \mathfrak{S}^*} |D_n(J, X)|$$

où \mathfrak{S}^* est l'ensemble des parties de $[0, 1]^d$ de la forme $\prod_{i=1}^d [0, b_i]$.

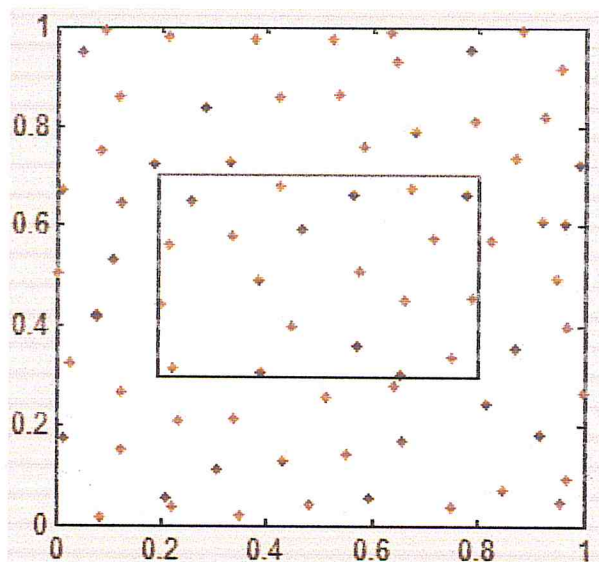


Figure 2.2 Les 80 premiers points en dimension 2 avec un sous-ensemble J défini par x et x' pour le calcul de la discrédance extrême

Pour $x = (0.2, 0.3)$ et $x' = (0.8, 0.7)$, le volume du rectangle J est égal à 0.24, et la proportion de points dans J est $18/80 = 0.225$. On a donc $D_n(J, X) = 0.015$. En effectuant le même procédé sur plusieurs sous-ensembles J et en prenant la valeur maximale acquise, on obtient la discrédance extrême de la suite X .

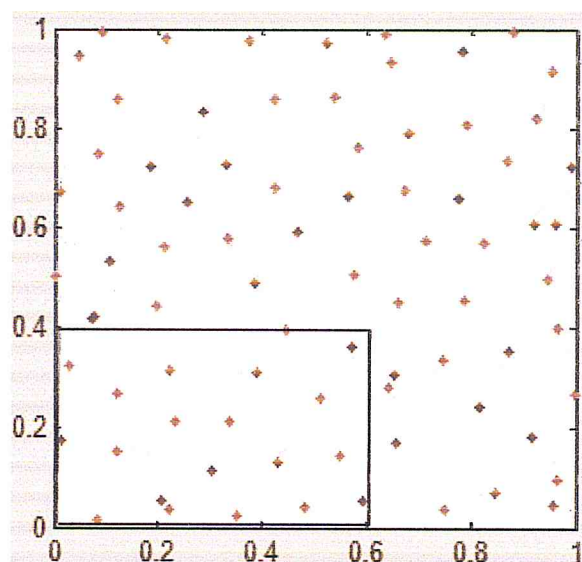


Figure 2.3. Un sous-ensemble J pour le calcul de la discr ance l'origine

$D_n(X)$ et $D_n^*(X)$ sont des discr ances en norme L_∞

2.2.2 Discr ance en norme 2

Seule la discr ance en norme L_2 est facilement calculable quelle que soit la dimension. En voici plusieurs. On note X la suite de n points x^1, \dots, x^n sur $[0,1]^d$. Cette section pr sente dans un premier temps la d finition des diff rentes discr ances en norme 2 ainsi que leurs approches calculatoires.

2.2.2.1 Discr ance extr me et   l'origine

D finition 1:

La discr ance L_2 d'une suite de n points x^1, \dots, x^n sur $[0,1]^d$ est d finie par

$$DL2_n(X)^2 = \int_{[0,1]^{2d}} (D_n(J.X))^2 dadb$$

où J est l'ensemble des parties de $[0.1]^d$ de la forme $\prod_{i=1}^d [a_i, b_i]$.

Définition2:

La discrédance L_2 à l'origine d'une suite de n points x^1, \dots, x^n sur $[0.1]^d$ est définie par

$$DL2_n^*(X)^2 = \int_{[0.1]^d} (D_n(J.X))^2 db$$

où J est l'ensemble des parties de $[0.1]^d$ de la forme $\prod_{i=1}^d [0, b_i]$.

En dimension d , les calculs de $DL2_n$ et $DL2_n^*$ peuvent se faire par les formules explicites suivantes :

$$DL2_n(X)^2 = \left(\frac{1}{12}\right)^d - \frac{2^{1-d}}{n} \sum_{i=1}^n \prod_{j=1}^d x_j^i (1 - x_j^i) + \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^n \prod_{j=1}^d (1 - \max(x_j^i - x_j^k)) \min(x_j^i - x_j^k)$$

$$DL2_n^*(X)^2 = \left(\frac{1}{3}\right)^d - \frac{2^{1-d}}{n} \sum_{i=1}^n \prod_{j=1}^d (1 - (x_j^i)^2) + \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^n \prod_{j=1}^d (1 - \max(x_j^i - x_j^k))$$

2.2.3 Exemples

Quelle que soit la discrédance, plus la valeur calculée pour un plan donné est faible et plus les points du plan seront distribués uniformément, à 49 points en dimension 2.

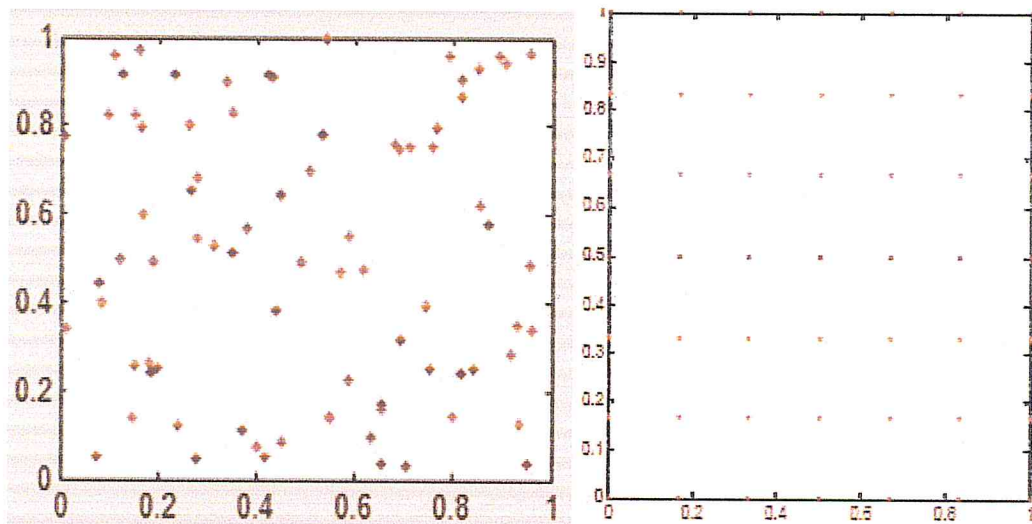


Figure 2.4. De la gauche vers la droite. Un plan alatoire, et un plan factoriel complet 49 points en dimension 2

Le tableau ci-dessous montre que pour le plan factoriel, on obtient les discrèpances les plus mauvaises. Ceci implique que la discrèpance seule ne suffit pas à mesurer la qualité d'un plan car nous ne souhaitons pas uniquement mesurer l'uniformité des points mais aussi le remplissage de l'espace. A noter que cette remarque est valable pour n'importe quel nombre de points. [9]

	Aléatoire	Factoriel
N=16	0.26	0.44
N=25	0.21	0.36
N=36	0.26	0.31
N=49	0.12	0.27

Tableau 2.1. Discrèpances en norme infinie pour $n=16, 25, 36$ et 49 pour les 2 types de plans suivants : aléatoires, plan factoriel en dimension 2

2.3 Critères de distance

L'objectif des critères basés sur la distance entre les points est de quantifier la proximité entre une distribution de points donnée et celle d'une grille régulière. Nous avons choisi de discuter uniquement des critères d'uniformité les plus courants basés sur la distance entre les points voisins.

L'idée ici est de générer des plans dont les points sont proches d'une grille régulière mais pas exactement ceux d'une grille à cause des mauvais alignements que cela peut engendrer. On rappelle que nous recherchons des plans avec des distributions quasi-périodiques c'est-à-dire qui sont un bon compromis entre une grille régulière et une bonne uniformité.

On note $\text{dist}(x^i, x^k)$ la distance euclidienne entre deux points x^i et x^k telle que :

$$\text{dist}(x^i, x^k) = \left[\sum_{j=1}^d (x_j^i - x_j^k)^2 \right]^{1/2}$$

2.3.1 Mesure de recouvrement

Définition:

Soit $X = \{x^1, x^2, \dots, x^n\} \subset [0,1]^d$ une suite de n points en d dimensions.

On définit la mesure de recouvrement λ par

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

avec $\gamma_i = \min_{i \neq k} \text{dist}(x^i, x^k)$ désignant la distance minimale entre le point x^i et les autres points de la suite et $\bar{\gamma} = \frac{1}{n} \sum_{i=1}^n \gamma_i$.

où $\bar{\gamma}$ la moyenne des γ_i

Si les points sont sur une grille régulière, on a $\gamma_1 = \dots = \gamma_n = \bar{\gamma}$ et donc $\lambda = 0$.

Ainsi, plus λ est faible, et plus les points sont proches de ceux d'une grille régulière.

2.3.2 Le rapport de distances

Définition:

Soit $X = \{x^1, x^2, \dots, x^n\} \subset [0,1]^d$ une distribution de n points en d dimensions.

On définit le rapport de distance par

$$R = \frac{\max_{i=1, \dots, n} \gamma_i}{\min_{i=1, \dots, n} \gamma_i}$$

avec $\gamma_i = \min_{i \neq k} \text{dist}(x^i, x^k)$.

Dans le cas où les points sont sur une grille régulière, on a $\gamma_1 = \dots = \gamma_n$ et donc $R = 1$.

Ainsi, plus R est proche de 1, et plus les points sont proches de ceux d'une grille régulière.

2.3.3 Distances Maximin et Minimax

Définition:

On définit les critères suivants au sens de la distance euclidienn

$$\min Dist = \min (\min_{i \neq k} \text{dist}(x^i, x^k))$$

et

$$AvgDist = \frac{1}{n} \sum_{x_i \in X} \min_{x_k \in X} \text{dist}(x^i, x^k) \quad i \neq k$$

où $X = \{x^1, \dots, x^n\}$ est le plan d'expériences à n points en d dimensions.

Maximiser le critère $\min Dist$ revient à considérer le critère de distance **maximin**.

Si la valeur de $\min Dist$ est grand, alors il n'y a aucun point qui soit proche au sens de cette grandeur. Ainsi, plus les critères $\min Dist$ et $AvgDist$ seront grands et plus cela traduira une distribution uniforme.

Définition:

On définit les critères suivants au sens de la distance euclidienn

$$\max Dist(D, X) = \max_{x_i \in D, x_k \in X} (\max_{i \neq k} dist(x^i, x^k))$$

et

$$EDist = \int_D \min_{x \in X} dist(x^i, x^k) dx^i$$

où D est un ensemble de points du domaine expérimental et $X = \{x^1, \dots, x^n\}$ le plan d'expériences à n points en dimension d .

Minimiser le critère $\max Dist$ revient à considérer le critère de distance **minimax**. Le principe du critère minimax est de mesurer la distance maximale entre un point du domaine expérimental n'appartenant pas au plan, qui varie, et les points du plan d'expériences. Afin de calculer les critères, on prend généralement pour les points de D une grille régulière. Aussi, en grande dimension il n'est plus envisageable de considérer ce critère sans prendre pour l'ensemble D un plan aléatoire.

Le critère $EDist$ correspond au critère de dispersion

2.3.4 Exemple

Considérons à présent la valeur des critères basés sur les mesures de distance pour les différents plans de l'exemple précédent (voir Tableau 2.1 et Figure 2.4)

N	Aléatoire				Factoriel			
	λ	R	MinDist	MaxDist	λ	R	MinDist	MaxDist
16	0.5	7.56	0.049	0.43	$3 \cdot 10^{-3}$	1	0.33	0.23
25	0.43	4.03	0.049	0.32	0	1	0.25	0.17
36	0.68	19.8	0.012	0.29	$3 \cdot 10^{-3}$	1	0.2	0.14
49	0.58	11.14	0.016	0.19	$2 \cdot 10^{-3}$	1	0.17	0.11

Tableau 2.2. Valeurs des critères de distance pour les 2 types de plans suivants : alatoires, plan factoriel où $n = 16, 25, 36, 49$

L'ensemble D considéré pour calculer le critère **MaxDist** est un plan factoriel, ce qui donne dans ce cas une bonne interprétabilité à ce critère. A noter que l'ensemble D doit être formé de beaucoup plus de points que l'ensemble de points testés pour rendre compte du positionnement des points et que ces points ne peuvent pas être simplement choisis de manière aléatoire.

Conclusions. Il est nécessaire de considérer les deux catégories de critères - discrédance et distance - pour déterminer si la distribution d'un plan remplit bien l'espace ou non. Cependant, ce sont surtout des critères de comparaison et non de qualification.

2.3.5 Algorithmes d'amélioration du critère de distance

Le principe de base de ces algorithmes consiste à déterminer un sous-ensemble de points, à partir d'un grand ensemble de candidats, qui améliore un critère géométrique de remplissage de l'espace.

2.3.5.1 Plans minimax

Dans ce type de procédé, les points de l'ensemble candidat ne servent pas seulement pour les points du plan mais permettent aussi de déterminer le critère de remplissage de l'espace afin d'évaluer la qualité du plan. Le critère utilisé ici cherche à minimiser la distance maximale entre les points de l'ensemble candidat et les points du plan.

Afin de trouver le plan optimal au sens de ce critère, l'algorithme utilise une distribution initiale aléatoire et fait diminuer le critère à l'aide d'un simple algorithme d'échange du même type que celui utilisé pour construire les plans à faible discrédance. Cependant, ici, les échanges se font entre les points du plan et un ensemble, de points candidats, fini qui sert à calculer le critère que l'on cherche à optimiser.

Le principe de l'algorithme est simple : pour un point donné du plan, on remplace ce point avec un de l'ensemble candidat et on examine si cet échange permet d'améliorer le critère. Si l'amélioration est avérée, le nouveau point est adopté dans le plan et l'ancien point est ajouté à l'ensemble des candidats. Ce processus continue jusqu'à ce qu'il n'y ait plus d'échange possible ou que le nombre d'itérations préalablement fixé soit atteint. Ainsi la convergence de l'algorithme est assurée, mais aucune garantie n'est donnée d'avoir généré le meilleur plan possible.

Algorithme 1. Algorithme pour générer un plan minimax par**Le rapport de distances**

Initialisation : choisir un plan d'expériences x^1, \dots, x^n selon une loi de probabilité donnée (loi uniforme par exemple)

POUR $n = 1$ à N

Choisir une expérience x^i au hasard

Simuler y^i uniformément dans $[0.1]^d$

Accepter le changement $x^i = y^i$ si

$$|R_k - 1| \succ |R_{k-1} - 1|$$

où $y = (x^1, \dots, x^{i-1}, y^i, x^{i+1}, \dots, x^n)$

Fin pour N

2.4 Critère d'entropie

Ce critère est différent de ceux présentés précédemment dans le sens où il ne permet pas à proprement parler d'étudier l'uniformité et le remplissage de l'espace d'un plan en phase exploratoire. En effet, le calcul de l'entropie n'est bien souvent envisageable que si la distribution sous-jacente est connue ; hypothèse que nous ne possédons généralement pas en phase exploratoire.

2.4.1 Définition de l'entropie

L'entropie d'une variable aléatoire continue X à densité de probabilité f est définie par :

$$H(X) = - \int_{x \in R} f(x) \log(f(x)) dx$$

Par ailleurs, l'entropie d'un vecteur aléatoire $X = (X_1, \dots, X_d)$ à valeurs dans \mathbb{R}^d et de même densité f se définit par :

$$H(X) = - \int_{x \in \mathbb{R}^d} f(x) \log(f(x)) d\mu(x)$$

où f est la densité de probabilité et μ la mesure de Lebesgue.

2.4.2 Entropie et quantité d'information.

L'entropie est une quantité d'information. Afin de comprendre cette notion, nous allons considérer une variable aléatoire discrète X dont la distribution est donnée par $p(X = x_i) = p_i$. L'entropie de X est alors définie par :

$$H(X) = - \sum_{i=1}^n p_i \log(p_i)$$

D'une part, on remarque que si toutes les valeurs de X sont distribuées de manière uniforme (c'est-à-dire où $p_1 = p_2 = \dots = p_n$) alors X est la plus imprévisible possible et $H(X) = \log(n)$ atteint sa valeur maximale. Si au contraire X prend une seule valeur avec la probabilité 1 alors X est complètement prévisible et l'entropie associée sera minimale avec $H(X) = 0$.

Cet exemple montre que la quantité d'information disponible sur la distribution peut se mesurer par $I = -H$. Cette quantité d'information I apportée par un paramètre θ est la seule fonction ayant la propriété d'additivité à une constante près et la propriété de continuité.

2.4.3 Plans à entropie maximale

La définition générale des plans à entropie maximale nécessite d'avoir la réponse aux points du plan d'expériences.

La méthode qui suit propose un remplissage de l'espace, en distribuant des points selon une matrice de corrélation spatiale. Dans le cas particulier d'un processus gaussien centré f , Ainsi, générer un plan à entropie maximale à, maximiser le déterminant de la matrice de corrélation.

Cette équivalence existe uniquement si les réponses aux points suivent une loi normale multidimensionnelle, mais sans aucune condition sur la structure de covariance.

Soit $X = (X^1, X^2, \dots, X^n)$ un vecteur de variables aléatoires, la matrice de variance-covariance de X s'écrit :

$$C(X) = \begin{pmatrix} \sigma_1^2 & \text{cov}(X^1 X^2) & \text{cov}(X^1 X^n) \\ & \sigma_2^2 & \\ & & \sigma_n^2 \end{pmatrix}$$

Si les variables X_i sont réduites, $C(X)$ s'identifie avec la matrice de corrélation :

$$C(X) = \sigma^2 \begin{pmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{pmatrix}$$

où ρ_{ij} est une fonction de la distance entre les points i et j , Le déterminant de $C(X)$ est maximum pour ρ_{ij} , c'est-à-dire lorsque les points pris deux à deux sont éloignés l'un de l'autre d'une distance supérieure à la portée a de la fonction de corrélation spatiale. [9]

On cherche donc à maximiser le déterminant de $C(X)$ par un algorithme

Algorithme 2. Algorithme pour générer un plan à entropie maximale

Initialisation : choisir un vecteur de n points $X_{(0)}$ au hasard

Calcul du déterminant de la matrice de corrélation spatiale associée : $\det(C(X_0))$

POUR $k = 1$ à N_{Max}

 Choisir une expérience i au hasard dans $\{1, \dots, n\}$

 Simuler z^i uniformément dans $[0,1]^d$

 Accepter de remplacer x^i par z^i si $\det(C(X_k)) > \det(C(X_{k-1}))$

Fin pour N

CHAPITRE III

PLANS D'EXPÉRIENCES PAR MÉTHODES MCMC

3.1 Introduction

les méthodes usuelles de construction de plans d'expériences ne sont pas adaptées et sont généralement basées sur des heuristiques et non sur des fondements théoriques, nous introduisons un nouveau type de Space-Filling Designs spécifié à partir du processus ponctuel de Strauss. Ces plans construits à partir de méthodes Monte-Carlo par chaînes de Markov (McMC) sont facilement modulables et peuvent ainsi répondre à divers objectifs. Signalons en particulier les bonnes propriétés de projection des expériences sur les axes factoriels.

La première partie de ce chapitre argumentera l'introduction des Méthodes Monte Carlo et méthodes McMC en planification d'expériences numériques et présentera l'algorithme général de Métropolis- Hastings (MH). La deuxième partie sera consacrée aux chaînes de Markov afin d'étudier la convergence de l'algorithme de MH. Enfin, on présentera la nouvelle famille de plans en question en s'appuyant sur la méthodologie McMC décrite auparavant.

3.2 Méthodes Monte Carlo

Le terme **méthode de Monte-Carlo**, désigne une famille de méthodes algorithmiques visant à calculer une valeur numérique approchée en utilisant des **procédés aléatoires**, c'est-à-dire des techniques probabilistes.

Le nom de ces méthodes, qui fait allusion aux **jeux de hasard** pratiqués à **Monte-Carlo**, a été inventé en 1947 par **Nicholas Metropolis**, et publié pour la première fois en 1949 dans un article coécrit avec **Stanislaw Ulam**.

Les méthodes de Monte-Carlo sont particulièrement utilisées pour calculer des intégrales en dimensions plus grandes que 1 (en particulier, pour calculer des surfaces et des volumes). Elles sont également couramment utilisées en **physique des particules**, où des **simulations** probabilistes permettent d'estimer la forme d'un signal ou la sensibilité d'un détecteur. La comparaison des données mesurées à ces simulations peut permettre de mettre en évidence des caractéristiques inattendues, par exemple de nouvelles particules. [8]

La méthode de Monte Carlo se généralise facilement à d'autres problèmes que des calculs de volume. Supposons par exemple donné un espace probabilisé $(\Omega, \mathcal{F}, \mu)$, et une variable aléatoire $Y : \Omega \rightarrow \mathbb{R}$. On voudrait estimer l'espérance de Y . Pour cela, l'algorithme de Monte Carlo consiste à générer des variables aléatoires indépendantes

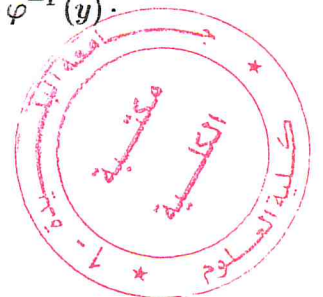
Y_1, Y_2, \dots, Y_n toutes de loi uniforme, et de calculer leur moyenne. Cette moyenne doit converger vers l'espérance cherchée.

Exemple (Cas unidimensionnelle). Une variable aléatoire unidimensionnelle Y s'obtient facilement à partir d'une variable de loi uniforme. Soit en effet U une variable uniforme sur $[0,1]$. Sa fonction de répartition est donnée par

$$F_U(u) = p(U \leq u) = \begin{cases} 0 & \text{si } u < 0 \\ u & \text{si } 0 \leq u \leq 1 \\ 1 & \text{si } u > 1 \end{cases}$$

On cherche une fonction φ telle que la variable $Y = \varphi(U)$ admette la fonction de répartition prescrite $F_Y(y)$. Or on a

$$F_Y(y) = P(Y \leq y) = P(\varphi(U) \leq y) = P(U \leq \varphi^{-1}(y)) = \varphi^{-1}(y).$$



Il suffit donc de prendre $Y = F_Y^{-1}(U)$.

Par exemple, pour générer une variable de loi exponentielle, dont la fonction de répartition est donnée par $F_Y(y) = 1 - \exp(-\lambda y)$, il suffit de prendre

$$Y = -\frac{1}{\lambda} \log(1 - U)$$

3.3 Méthodes McMC

On introduit dans ce chapitre des plans qui sont par nature aléatoires. Ils sont définis par une loi de probabilité sur le domaine cubique $[0, 1]^d \times \dots \times [0, 1]^d$:

$$\pi(dx) = k \times \pi(x) dx ; x = (x^1, \dots, x^n) \quad (1)$$

Le choix de la densité $\pi(x)$ sera lié aux contraintes imposées aux plans d'expériences. Nous verrons ultérieurement les raisons pour lesquelles on considère un processus de Strauss ou plus généralement un processus ponctuel de Gibbs.

A ce stade, disons simplement qu'il est impossible en pratique de simuler directement de tels plans à cause de la relative complexité de la loi $\pi(dx)$. En particulier, il n'existe pas d'expression analytique simple de la constante de normalisation k intervenant dans la densité du processus (1). Comme l'espace est de grande dimension ($n \times d$), l'estimation de cette constante n'est pas envisageable.

L'alternative classique est de simuler une chaîne de Markov de loi asymptotique (ou stationnaire) $\pi(\cdot)$. C'est le principe des méthodes dites de Monte-Carlo par chaînes de Markov (McMC). La souplesse d'une telle approche réside dans le fait que l'on approche la loi π par itérations successives. Le choix du mécanisme de transition de la chaîne de Markov rend une transition d'état à état facile à simuler. Les processus que l'on souhaite simuler ne sont connus que par leur densité non normalisée, ainsi la construction de la chaîne de Markov ne doit pas faire intervenir la constante de normalisation. C'est le cas de l'algorithme de Metropolis-Hastings (MH) qui consiste à

proposer à chaque étape des transitions de l'état courant x vers un état y quelconque. La transition proposée est en suite acceptée ou non selon une certaine probabilité. Le taux d'acceptation fait intervenir la densité de manière à assurer la convergence de l'algorithme vers la loi π . [9]

L'algorithme de MH se décrit donc en 2 étapes :

- on fait une proposition de changement d'états x à y selon une loi de probabilité $Q(x, dy)$
- on accepte y avec la probabilité $a(x, y)$; sinon, on reste dans l'état x .

$$a(x, y) = \min \left(1, \frac{\pi(y)}{\pi(x)} \right)$$

qui favorise les configurations les plus probables pour la loi π . On remarque que le calcul de la probabilité $a(x, y)$ ne fait pas intervenir la constante de normalisation k .

Remarque. Le choix de l'état de destination y ne dépend pas forcément de la densité à simuler.

3.4 Chaînes de Markov

On appellera chaîne de Markov toute suite $(X_p)_{p \geq 0}$ de variables aléatoires à valeurs dans un même espace E (désigné espace des états ou configurations) et obéissant à une dynamique de la forme :

$$X_0 \text{ de loi initiale } \nu \text{ quelconque ; } X_{p+1} = f(X_p, U_p) \text{ pour } p \geq 0$$

où les U_p sont des vecteurs aléatoires indépendants et identiquement distribués.

La fonction f , qui est à valeurs dans E , représente typiquement une itération ou boucle d'un algorithme de simulation MCMC. La taille des vecteurs U_p est le nombre d'appels à la fonction *RAND* dans cette boucle (cette taille peut être finie ou non).

La traduction probabiliste de la dynamique ci-dessus est qu'à tout instant $p \geq 0$, la loi conditionnelle de X_{p+1} sachant X_p, X_{p-1}, \dots, X_0 (présent + passé) est égale à la loi conditionnelle de X_{p+1} sachant X_p . Dit autrement, pour prévoir l'évolution

future de la suite à un instant p , connaissant la dernière valeur X_p , le passé X_{p-1}, \dots, X_0 n'apporte aucune information supplémentaire. Cela ne signifie pas que l'évolution future soit indépendante du passé mais qu'elle n'en dépend qu'à travers le présent X_p : le futur ne dépend du passé qu'à travers le présent! On reconnaît ici précisément la définition usuelle d'une chaîne de Markov à partir de cette propriété dite de Markov (simple).

La question centrale qui va se poser est de savoir si partant d'une loi initiale ν arbitraire (celle de X_0), la loi obtenue au bout de p itérations (donc la loi de X_p) approche la loi stationnaire π de la chaîne (à savoir la loi que l'on cherche à simuler). Pour aborder cette question, on se placera dans un premier temps dans le cadre classique des chaînes de Markov à espace d'états discrets. C'est l'objet du paragraphe suivant qui considère un cas d'école dans un but essentiellement pédagogique. [6]

3.4.1 Un cas simple dans un espace d'états discret

Les chaînes de Markov les plus simples sont celles dont les états ne prennent qu'un nombre fini ou dénombrable de valeurs. Cette section est consacrée à une présentation de ces chaînes élémentaires dans le cadre de la planification d'expériences. En effet, nous allons mettre en oeuvre la méthodologie McMC sur un exemple simplifié de planification d'expériences de manière à se familiariser avec les principaux concepts relatifs aux chaînes de Markov.

Pour cela, les plans d'expériences qui nous intéressent sont composés seulement de $n = 2$ expériences distribuées dans l'intervalle $[0,1]$ (cas de la dimension $d = 1$). Ainsi, un plan d'expériences x est un couple de points (x^1, x^2) dans $[0,1]^2$. L'espace E est celui des états de l'algorithme de construction itérative d'un plan ou de la chaîne de Markov.

Pour éliminer la difficulté essentiellement technique liée au fait que E est un espace d'états continu, on va encore radicalement le simplifier en le réduisant aux 4 plans

d'expériences possibles.

Avec la numérotation ci-dessus, on convient que $E = \{1, 2, 3, 4\}$ ce qui fait de E un espace discret fini. On confond tout état i de avec le plan (x^1, x^2) qui lui correspond.

Une représentation des éléments de E dans l'espace euclidien $[0, 1]^2$ des plans d'expériences est la suivante :

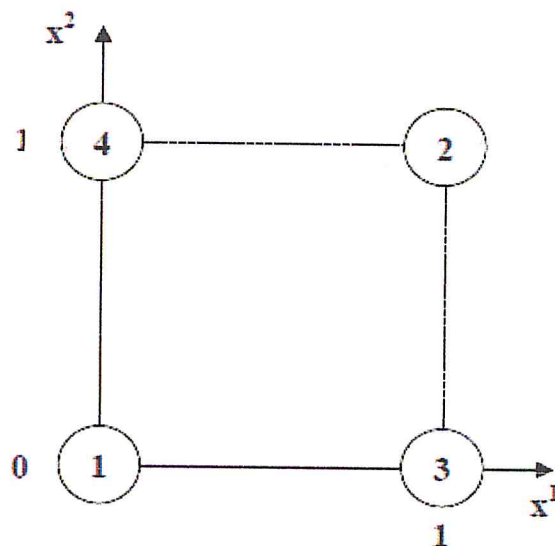


Figure (3.1): Représentation des éléments de E dans $[0, 1]^2$

Soit maintenant π la distribution sur l'espace E des 4 plans d'expériences définie par

$$\pi = (\pi_1 \pi_2 \pi_3 \pi_4) = k \times (\gamma \gamma 1 1)$$

avec π_i la probabilité de l'état i et où $0 < \gamma < 1$ est assimilé à un coefficient d'interaction entre les 2 expériences. Enfin, $k = \frac{1}{2 \times (1 + \gamma)}$ est la constante de normalisation.

Ainsi, on donne une probabilité plus forte aux états 3 et 4 et plus faible aux états 1 et 2 et Pour simuler π , on propose l'algorithme de Métropolis-Hastings suivant

Algorithme 3. Algorithme de Métropolis-Hastings

Initialisation : Soit $n = 2$. Choisir $x = (x^1, x^2)$, par exemple selon la loi

$$\text{uniforme } v = \left(\frac{1}{4} \frac{1}{4} \frac{1}{4} \frac{1}{4}\right)$$

POUR N = 1 à NMC

POUR k = 1 à n

Choisir i uniformément dans $\{1, 2\}$ ($i = 1$ ou 2 avec probabilité $1/2$)

Simuler u uniformément dans $\{0, 1\}$ ($u = 0$ ou 1 avec probabilité $1/2$)

Accepter le changement $x^i = u$ avec la probabilité

$$\min\left(1, \frac{\pi(y)}{\pi(x)}\right)$$

où $y = (u, x^2)$ si $i = 1$ et $y = (x^1, u)$ si $i = 2$

Fin pour k

Fin pour N

On notera qu'il s'agit d'un algorithme itératif, NMC correspondant au nombre d'itérations McMC effectuées. Notons P la matrice de transition correspondant à chaque itération élémentaire dans la boucle « Pour $k = 1$ à n » et $k = P^n$ la matrice de transition correspondant à l'ensemble du balayage. Ainsi, la suite $X_{(0)}, X_{(1)}, \dots, X_{(NMC)}$ des plans générés par cet algorithme est la réalisation d'une chaîne de Markov $(X_P)_{P \geq 0}$ de probabilité la matrice p .

$$P = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{\gamma}{4} & \frac{\gamma}{4} & 1 - \frac{\gamma}{2} & 0 \\ \frac{\gamma}{4} & \frac{\gamma}{4} & 0 & 1 - \frac{\gamma}{2} \end{pmatrix}$$

où $p(i, j)$ s'interprète comme la probabilité pour passer du plan d'expériences i au plan j est égale au produit des probabilités:

- de choisir le point x^1 (parmi 2 possibles) du plan d'expériences à savoir $i = 1$
- de choisir la direction (parmi 2 possibles) pour se retrouver dans l'état 1 à savoir $u = 1$
- d'accepter l'état j venant de l'état i

On a donc,

$$p(i, j) = \frac{1}{2} (\text{choix du point}) * \frac{1}{2} (\text{choix de la direction}) * (\text{probabilité d'accepter}) .$$

3.4.2 Réversibilité et mesure invariante

Les transitions π -réversibles vont jouer un rôle important dans la construction des algorithmes de simulation. On dit que le mécanisme de transition P sur E est π -réversible s'il vérifie la condition dite de réversibilité

$$\pi_i P(i, j) = \pi_j P(j, i)$$

On vérifie facilement que c'est bien le cas ici et cela signifie que, sous la distribution π , la probabilité de passer de x à y est la même que de passer de y à x .

La première conséquence pour π est qu'elle est une loi invariante (ou stationnaire) pour P , c'est-à-dire que l'on a l'égalité suivante

$$\pi P = \pi$$

En effet, on a :

$$(\pi P)_i = \sum_{j=1}^n \pi_j P(j, i) = \pi_i \sum_{j=1}^n P(i, j) = \pi_i$$

3.5 *Processus ponctuels et planification d'expériences*

L'idée principale est de voir chaque expérience comme un point ou particule chargée de l'espace $\Omega = [0, 1]^d$. On assimile ainsi les n expériences à la réalisation (conditionnelle au nombre n de points) d'un processus ponctuel de manière à bien occuper l'espace, il a été naturel de considérer le processus de Strauss (1975) qui permet de rendre compte de répulsion entre particules. Ainsi, les n expériences sont vues comme la photographie à un instant donné de n particules de même charge électrique en mouvement dans une boîte d -dimensionnelle.

Pour générer un tel processus, nous utiliserons les techniques usuelles de simulation par chaînes de Markov et plus précisément l'algorithme de Metropolis-Hastings (M; H). C'est essentiellement la propriété de Markov spatiale du processus ponctuel de Strauss qui en pratique permet de simuler de manière efficace ces nouveaux plans d'expériences, notamment en grande dimension. Pour l'étude de tels plans (aléatoires), le calcul de la moyenne des critères d'uniformité usuels est envisagé.

C'est la dynamique markovienne de l'algorithme de MH et ses bonnes propriétés de mélange ou d'ergodicité qui vont permettre ce calcul. Ainsi, on peut parler de planification d'expériences par méthodes de Monte-Carlo par chaînes de Markov (McMC).

[9]

3.5.1 Le processus ponctuel de Strauss

On considère le processus de Strauss avec interaction de type répulsion dans le domaine $\Omega = [0, 1]^d$. La loi $\pi(\cdot)$ de ce processus conditionnellement au nombre n de points est donnée par :

$$\pi(x) = k \times \gamma^{s(x)}$$

où $0 < \gamma < 1$ est un coefficient de répulsion, k la constante de normalisation, et où la fonction $s(x)$ assimilée à un potentiel global d'énergie est définie par :

$$s(x) = s(x^1 \dots x^n) = \sum_{i < j} 1_{\|x^i - x^j\| \leq R}$$

Il s'agit simplement du nombre de paires d'expériences $\{x^i, x^j\}$ telles que la distance entre x^i et x^j est inférieure ou égale à R .

En termes de particules chargées, c'est aussi le nombre global d'interactions limité uniquement aux interactions de paires (interactions dites d'ordre 2). Chaque particule a une sphère d'influence de rayon $R/2$, l'interaction se produisant exactement lorsque 2 sphères quelconques se rencontrent.

Le choix $0 < \gamma < 1$ permet de donner une probabilité plus ou moins forte aux plans pour lesquels les expériences sont réparties de manière à interagir plus ou moins les unes par rapport aux autres. Le plan est d'autant plus vraisemblable que le nombre d'interactions $s(x)$ est faible. Nous ne considérons pas ici le cas limite où $\gamma = 0$, qui interdirait toute paire de points de distance inférieure ou égale à R . Le cas $\gamma = 1$ correspond exactement à l'indépendance (distribution conjointe uniforme).

Dans toute la suite, on se limitera à des réalisations de processus ponctuels conditionnellement à un nombre n de points spécifiés à l'avance. Ainsi, dans la dynamique de simulation, ce nombre n n'aura pas la possibilité d'évoluer.

La non-existence d'expression analytique explicite pour la constante de normalisation k intervenant dans la densité du processus (1), même pour des tailles modestes de n et d , motive l'utilisation des méthodes *McMC*. Dans le contexte de la planification d'expériences, un état de la chaîne est un plan de n expériences.

L'algorithme de MH va toujours procéder en 2 étapes :

- une phase de proposition de changement d'état où le nouvel état ne diffère que par une seule expérience
- et une autre d'acceptation ou de rejet de ce changement.

Algorithme 4. Algorithme de Metropolis-Hastings pour générer des plans d'expériences dits de Strauss

Initialisation : choisir un plan d'expériences x^1, \dots, x^n selon une loi de probabilité donnée (loi uniforme par exemple)

POUR N = 1 à NMC

POUR k = 1 à n

Choisir une expérience x^i au hasard

Simuler y^i uniformément dans $[0,1]^d$

Accepter le changement $x^i = y^i$ avec la probabilité

$$a(x, y) = \min\left(1, \frac{\pi(y)}{\pi(x)}\right)$$

où $y = (x^1, \dots, x^{i-1}, y^i, x^{i+1}, \dots, x^n)$

Fin pour k

Fin pour N

La force de cet algorithme tient au fait que le calcul du rapport $\frac{\pi(y)}{\pi(x)}$ se limite à chaque fois à calculer

$$s(x^i) = \sum_{j \neq i} 1_{\|x^i - x^j\| \leq R} \text{ et } s(y^i) = \sum_{j \neq i} 1_{\|y^i - x^j\| \leq R}.$$

En effet, on vérifie que

$$\frac{\pi(y)}{\pi(x)} = \gamma^{s(y^i) - s(x^i)}$$

L'entier $s(x^i)$ est le nombre de voisins de l'expérience x^i , i.e. les points contenus dans la sphère centrée en x^i et de rayon R . Quant à $s(y^i)$, c'est le nombre de voisins de l'expérience i après la transition vers y^i .

On notera que ces plans sont bien de nature aléatoire comme par exemple les plans uniformes, Cependant, ils sont construits à partir d'une loi de probabilité π spécifiée au préalable. C'est à travers cette loi et ses extensions (processus de Gibbs de potentiel plus général) qu'il sera possible de traduire les propriétés souhaitées. L'utilisation de l'algorithme de MH pour la simulation de la loi π sera justifiée en vérifiant que les hypothèses du théorème énoncé au paragraphe (3.4.1) sont satisfaites pour le mécanisme de transition associé.

Désormais, nous parlerons de *plans de Strauss*.

3.5.2 Extensions : processus ponctuel de Gibbs

Plusieurs généralisations du processus ponctuel de Strauss sont envisageables. Par exemple, on peut faire intervenir les triplets de points distants deux à deux de moins de R , Nous nous contenterons ici seulement des interactions de paires qui nous paraissent suffisantes pour atteindre nos objectifs.

3.5.2.1 Plans de Strauss-Gibbs

Les plans de Strauss sont issus d'une loi de probabilité qui rend plus vraisemblable un nombre relativement faible d'interactions globales (ainsi que sur les marges, voir paragraphe 3.5.2.1). Pour pallier le phénomène d'agglomérats de points, notamment lorsque le rayon est inadapté, on prend en compte la distance réelle entre les points pour mesurer la force de l'interaction et non plus se contenter de savoir s'il y a interaction ou pas (sphères de rayon $R/2$ qui se rencontrent).

Cela nous amène à considérer la famille plus large des processus ponctuels dont la distribution π est de la forme :

$$\pi(x) = \alpha \exp(-U(x))$$

où

$$U(x) = \beta \sum_{i < j} \varphi(\|x^i - x^j\|) + \sum_{k=1}^d \beta_k \sum_{i < j} \varphi_k(\|x_k^i - x_k^j\|)$$

$$\beta = -\ln \gamma > 0; \quad \beta_k = -\ln \gamma_k > 0; \quad k = 1, \dots, d$$

et φ de $[0, \infty[$ dans R est une fonction continue décroissante telle que $\varphi(0) = 1$ et φ nulle à l'infini. On utilisera la famille ($\alpha > 0, R > 0$) de fonctions suivante

$$\varphi(r) = \left(1 - \frac{r}{R}\right)^\alpha \text{ si } 0 \leq r \leq R \quad (\varphi = 0 \text{ pour } r \geq R)$$

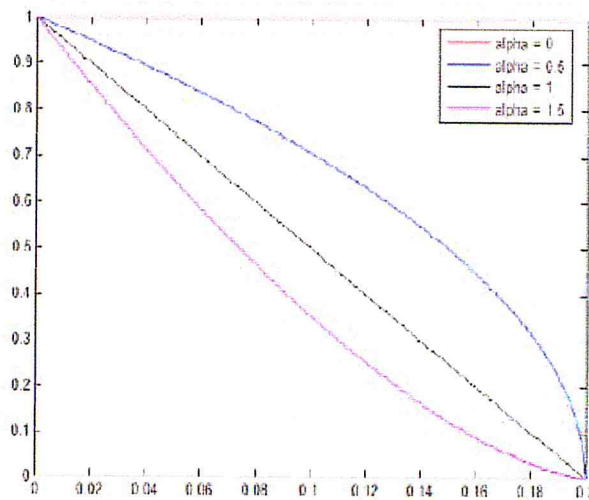


Figure (3.2): La fonction de potentiel puissance pour différents α et $R = 0.2$

Il s'agit de processus ponctuels de Gibbs de potentiel U , de potentiels locaux les fonctions φ et φ_k qui correspondent uniquement à des interactions de paires cas

$\alpha = 0$ (potentiel 0, 1 dit du « tout ou rien ») correspond exactement aux processus de Strauss discutés jusqu'ici, Cette extension réduit le risque de points agglomérés.

C'est la dernière extension envisagée dans cet article. L'algorithme de MH qui a été retenu est le suivant :

Algorithme 5. Algorithme de Metropolis-Hastings pour générer des plans d'expériences dits de Strauss_Gibbs

Initialisation : choisir un plan d'expériences x^1, \dots, x^n selon une loi de probabilité donnée (loi uniforme par exemple)

POUR N = 1 à NMC

POUR k = 1 à n

Choisir une expérience x^i au hasard

Simuler y^i uniformément dans $[0,1]^d$

Accepter le changement $x^i = y^i$ avec la probabilité

$$a(x, y) = \min(1, e^{U(x) - U(y)})$$

où $y = (x^1, \dots, x^{i-1}, y^i, x^{i+1}, \dots, x^n)$

Fin pour k

Fin pour N

Deux caractéristiques sont à la base de cet algorithme :

- la proposition de changement d'une expérience prise au hasard, notée $Q(x, dy)$
- la probabilité $a : E \times E \rightarrow [0, 1]$ d'accepter le changement $x \rightarrow y$ proposé

De manière précise, on a ici

$$(1) \quad Q(x, dy) = \sum_{i=1}^n \frac{1}{n} u(y^i) 1_{y^{(-i)}=x^{(-i)}} dy^i \text{ où } y^{(-i)} = x^{(-i)} \text{ signifie que } y_j = x_j \text{ pour } j \neq i$$

$$(2) \quad a(x, y) = \min(1, \exp(U(x) - U(y)))$$

CHAPITRE IV

IMPLÉMENTATION, DISCUSSIONS ET COMPARAISON

Les expérimentations ont été réalisées sur un PC équipé d'un processeur Intel(R) Core (TM) i3 -4005U CPU @ 1.70 GHz avec une mémoire de 4 GB sous Windows 7 Professionnel. Les algorithmes sont implémentés en **MATLAB**. Les programmes prennent comme donnée d'entrée des matrices d'expériences contenant des points et doivent être changées à la fin de l'algorithme pour arriver à l'optimalité.

Informations sur Matlab

Matlab est un langage simple et très efficace, optimisé pour le traitement des matrices, d'où son nom. Pour le calcul numérique, Matlab est beaucoup plus concis que les autres langages (C, Pascal, Fortran, Basic). On peut traiter la matrice comme une simple variable. Matlab contient également une interface graphique puissante, ainsi qu'une grande variété d'algorithmes scientifiques.

On peut enrichir Matlab en ajoutant des "boîtes à outils" (toolbox) qui sont des ensembles de fonctions supplémentaires, profilées pour des applications particulières (traitement de signaux, analyse statistique, optimisation, etc.).

4.1 Etude des plans de Strauss:

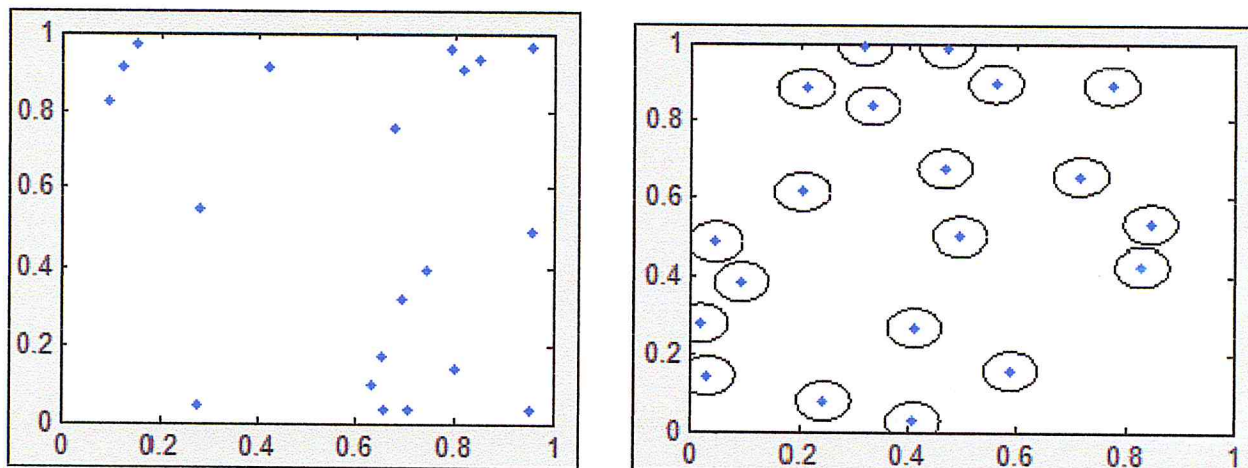
1^{ere} cas:

Un plan initial (aléatoire) de 20 points en dimension 2 et plan de Strauss obtenu à partir du plan initial avec la représentation des sphères de rayon $R/2$.

$$n = 20 \quad R = 0.1 \quad \gamma = 0.05 \quad NMC = 1000 \quad d = 2$$

$$\begin{array}{l}
 X_{initiale} = \left(\begin{array}{cc}
 0,814723 & 0,905791 \\
 0,126986 & 0,913375 \\
 0,632359 & 0,097540 \\
 0,278498 & 0,546881 \\
 0,957506 & 0,964888 \\
 0,157613 & 0,970592 \\
 0,957166 & 0,485375 \\
 0,800280 & 0,141886 \\
 0,421761 & 0,915735 \\
 0,792207 & 0,959492 \\
 0,655740 & 0,035711 \\
 0,849129 & 0,933993 \\
 0,678735 & 0,757740 \\
 0,743132 & 0,392227 \\
 0,655477 & 0,171186 \\
 0,706046 & 0,031832 \\
 0,276922 & 0,046171 \\
 0,0971317 & 0,823457 \\
 0,694828 & 0,317099 \\
 0,950222 & 0,034446
 \end{array} \right)
 \end{array}
 \quad
 \begin{array}{l}
 X_{finale} = \left(\begin{array}{cc}
 0.590271 & 0.158238 \\
 0.095973 & 0.385322 \\
 0.048225 & 0.485446 \\
 0.829966 & 0.425913 \\
 0.212208 & 0.884123 \\
 0.334969 & 0.838834 \\
 0.562447 & 0.896361 \\
 0.033658 & 0.142552 \\
 0.244650 & 0.079270 \\
 0.845443 & 0.533283 \\
 0.319596 & 0.994040 \\
 0.412655 & 0.269013 \\
 0.774835 & 0.891122 \\
 0.020358 & 0.278416 \\
 0.718705 & 0.655967 \\
 0.206457 & 0.614634 \\
 0.468765 & 0.675777 \\
 0.497321 & 0.501302 \\
 0.410515 & 0.026688 \\
 0.471727 & 0.988860
 \end{array} \right)
 \end{array}$$

$$X_{finale} = \begin{pmatrix} 0.590271 & 0.158238 \\ 0.095973 & 0.385322 \\ 0.048225 & 0.485446 \\ 0.829966 & 0.425913 \\ 0.212208 & 0.884123 \\ 0.334969 & 0.838834 \\ 0.562447 & 0.896361 \\ 0.033658 & 0.142552 \\ 0.244650 & 0.079270 \\ 0.845443 & 0.533283 \\ 0.319596 & 0.994040 \\ 0.412655 & 0.269013 \\ 0.774835 & 0.891122 \\ 0.020358 & 0.278416 \\ 0.718705 & 0.655967 \\ 0.206457 & 0.614634 \\ 0.468765 & 0.675777 \\ 0.497321 & 0.501302 \\ 0.410515 & 0.026688 \\ 0.471727 & 0.988860 \end{pmatrix} \quad X_{finale} = \begin{pmatrix} 0.015411 & 0.099880 \\ 0.597081 & 0.966888 \\ 0.290751 & 0.704957 \\ 0.604427 & 0.633771 \\ 0.973453 & 0.980299 \\ 0.004573 & 0.524788 \\ 0.471435 & 0.341724 \\ 0.314146 & 0.391338 \\ 0.827163 & 0.346051 \\ 0.037236 & 0.876893 \\ 0.009455 & 0.880579 \\ 0.689705 & 0.026992 \\ 0.057811 & 0.190556 \\ 0.931726 & 0.345345 \\ 0.682016 & 0.031289 \\ 0.989958 & 0.014062 \\ 0.976231 & 0.665360 \\ 0.640063 & 0.939529 \\ 0.659164 & 0.636091 \\ 0.331824 & 0.059379 \end{pmatrix}$$



Figure(4.1) :un plan initial (alatoire) et plan de Strauss obtenu partir du plan initial avec la representation des sphres de rayon $R/2$

Remarque. Sur les figures (4.1), on matérialisera les interactions entre expériences en traçant les cercles de rayon $R/2$, l'intersection de 2 cercles correspondant précisément à une interaction

2^{ere} cas:

Influence des paramètres

Les figures ci-dessous montrent l'influence du paramètre R sur la distribution finale. Le choix du rayon s'avère être important :

- un rayon trop petit engendre une distribution sans interaction mais avec de nombreuses lacunes (distribution uniforme),

- un rayon trop grand conduit à une distribution avec des agglomérats (clusters).

$$n = 20 \quad R = 0.05 \quad \gamma = 0.1 \quad NMC = 1000 \quad d = 2$$

$$n = 20 \quad R = 0.15 \quad \gamma = 0.1 \quad NMC = 1000 \quad d = 2$$

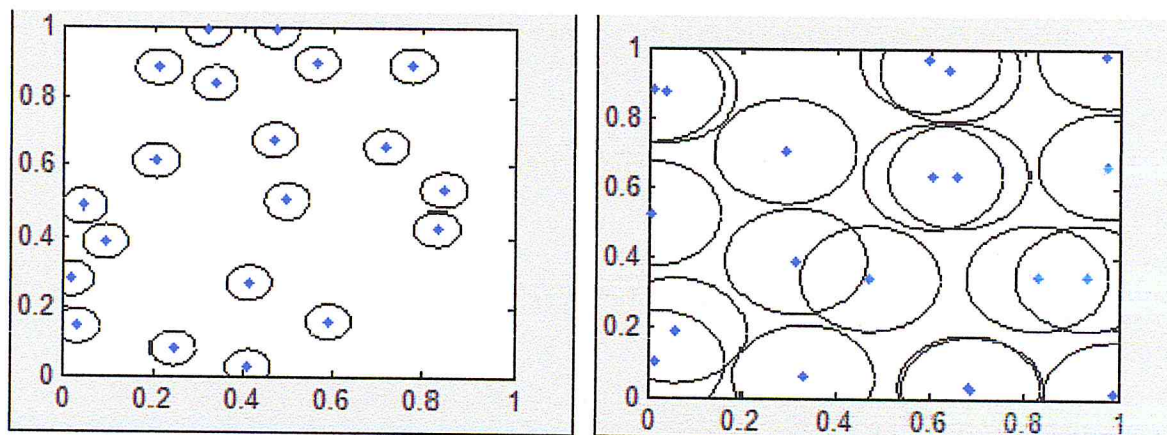


Figure (4.2) . A gauche, un plan de Strauss à 20 points en dimension 2 de paramètres $R = 0.1$ et $\gamma = 0.1$. A droite, un plan de Strauss de paramètres $R = 0.3$ et $\gamma = 0.1$.

3^{ere} cas:

La figure ci dessous permet d'illustrer et de comprendre, lorsque le rayon choisi est grand, que la loi π rend plus vraisemblable les agglomérats qui correspondent à un nombre global d'interactions compatible avec la loi π

$$n = 50 \quad R = 0.8 \quad \gamma = 0.1 \quad NMC = 1000 \quad d = 2$$

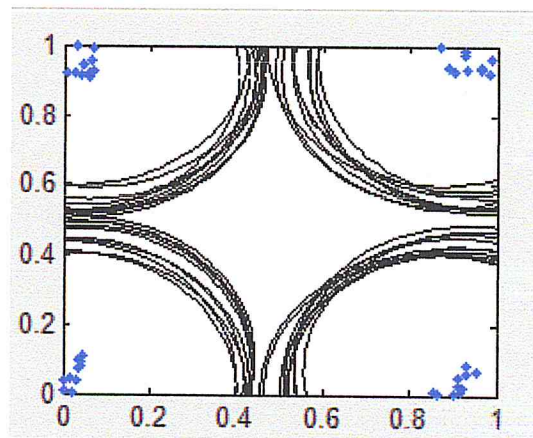


Figure (4.3) : Un plan de Strauss 50 points avec $R = 0.8$ et $\gamma = 0.1$

4^{ème} cas:

De même que pour le rayon d'interaction, il est important de bien fixer le paramètre de répulsion γ

$$n = 20 \quad R = 0.19 \quad \gamma = 0.5 \quad NMC = 1000 \quad d = 2$$

$$n = 20 \quad R = 0.19 \quad \gamma = 0.01 \quad NMC = 1000 \quad d = 2$$

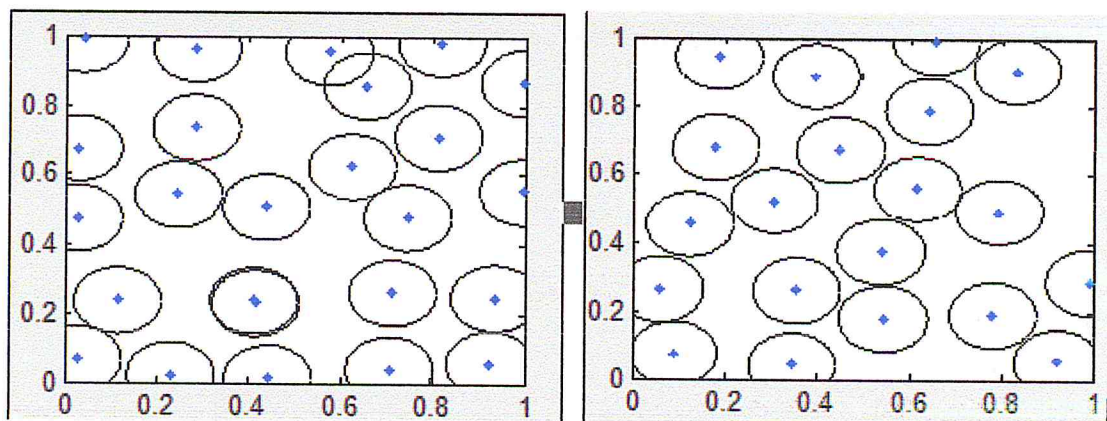


Figure (4.4) : A gauche, un plan de Strauss de paramtres $R = 0.19$ et $\gamma = 0.5$.

A droite, un plan de Strauss de paramtres $R = 0.19$ et $\gamma = 0.01$

La figure ci-dessus montre qu'il est plus aisé de générer une distribution répondant au critère de remplissage de l'espace avec un paramètre de répulsion faible. Cependant, un paramètre de répulsion trop faible contraint généralement les points à être alignés. Ce qui dégrade la répartition des points en projection, notamment sur les frontières du domaine expérimental .

Ainsi, il est délicat de bien fixer le paramètre de répulsion afin de trouver un compromis entre un remplissage de l'espace optimal et une répartition de points uniforme sur les marges de dimension 1. C'est pourquoi, nous proposons dans la suite une variante qui permet de corriger ce problème en considérant également les interactions de paires sur les d projections.

Le rayon d'interaction est le paramètre le plus sensible à régler. Pour un critère donné, la meilleure solution serait sans doute de tabuler cette valeur selon le nombre de points et la dimension du problème.

Une fois les problèmes de paramétrage résolus, les plans de Strauss permettent de bien répondre à la question du remplissage de l'espace en grande dimension.

4.2 Etude des plans de Strauss-Gibbs:

1^{ere} cas:

Cette extension réduit le risque de points agglomérés

$$n = 20 \quad R = 0.2 \quad \gamma = 0.1 \quad d = 2 \quad \gamma_1 = 0.1 \quad \gamma_2 = 0.1 \quad NMC = 100 \quad \alpha = 0$$

$$X_{initiale} = \begin{pmatrix} 0.814724 & 0.905792 \\ 0.126987 & 0.913376 \\ 0.632359 & 0.097540 \\ 0.278498 & 0.546882 \\ 0.957507 & 0.964889 \\ 0.157613 & 0.970593 \\ 0.957167 & 0.485376 \\ 0.800280 & 0.141886 \\ 0.421761 & 0.915736 \\ 0.792207 & 0.959492 \\ 0.655741 & 0.035712 \\ 0.849129 & 0.933993 \\ 0.678735 & 0.757740 \\ 0.743132 & 0.392227 \\ 0.655478 & 0.171187 \\ 0.706046 & 0.031833 \\ 0.276923 & 0.046171 \\ 0.097132 & 0.823458 \\ 0.694829 & 0.317099 \\ 0.950222 & 0.034446 \end{pmatrix} \quad X_{finale} = \begin{pmatrix} 0.445464 & 0.134685 \\ 0.499897 & 0.816779 \\ 0.503271 & 0.601971 \\ 0.716021 & 0.365197 \\ 0.644248 & 0.103249 \\ 0.242942 & 0.959343 \\ 0.987630 & 0.475466 \\ 0.002350 & 0.997748 \\ 0.914454 & 0.856733 \\ 0.245112 & 0.210091 \\ 0.287344 & 0.542641 \\ 0.295223 & 0.764906 \\ 0.043816 & 0.345848 \\ 0.905132 & 0.249481 \\ 0.718705 & 0.655967 \\ 0.690670 & 0.927137 \\ 0.081226 & 0.584645 \\ 0.856658 & 0.038155 \\ 0.029507 & 0.118186 \\ 0.423758 & 0.394988 \end{pmatrix}$$

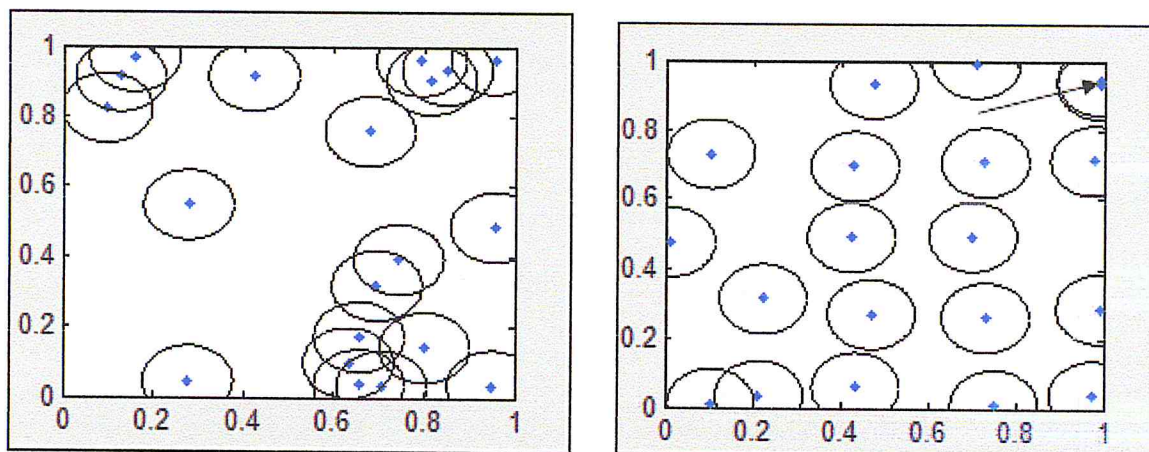


Figure (4.5) : A gauche. Un plan initiale. A droite. Un plan de Strauss obtenu
partir d'un potentiel tout ou rien

La Figure (4.5) permet de bien comprendre les interactions que l'on privilégie. En effet, sur le plan de droite généré avec un potentiel de type tout ou rien nous observons qu'il y a deux points de mêmes coordonnées (où le cercle est repéré par une flèche).

2^{ere} cas:

$n = 20$ $R = 0.82$ $\gamma = 0.1$ $d = 2$ $\gamma_1 = 0.1$ $\gamma_2 = 0.1$ $NMC = 1000$
 $\alpha = 0.5$

$$X_{initiale} = \begin{pmatrix} 0.814724 & 0.905792 \\ 0.126987 & 0.913376 \\ 0.632359 & 0.097540 \\ 0.278498 & 0.546882 \\ 0.957507 & 0.964889 \\ 0.157613 & 0.970593 \\ 0.957167 & 0.485376 \\ 0.800280 & 0.141886 \\ 0.421761 & 0.915736 \\ 0.792207 & 0.959492 \\ 0.655478 & 0.1711870 \\ 0.849129 & 0.933993 \\ 0.678735 & 0.757740 \\ 0.743132 & 0.392227 \\ 0.655478 & 0.171187 \\ 0.706046 & 0.031833 \\ 0.276923 & 0.046171 \\ 0.097132 & 0.823458 \\ 0.694829 & 0.317099 \\ 0.950222 & 0.034446 \end{pmatrix} \quad X_{finale} = \begin{pmatrix} 0.740803 & 0.997567 \\ 0.622386 & 0.688020 \\ 0.206069 & 0.213387 \\ 0.287454 & 0.442187 \\ 0.448020 & 0.569358 \\ 0.744693 & 0.188955 \\ 0.336071 & 0.984960 \\ 0.804266 & 0.784626 \\ 0.381558 & 0.765517 \\ 0.005257 & 0.261437 \\ 0.205238 & 0.632437 \\ 0.004818 & 0.690914 \\ 0.018495 & 0.958206 \\ 0.167810 & 0.834638 \\ 0.536434 & 0.906600 \\ 0.422979 & 0.197646 \\ 0.594999 & 0.330987 \\ 0.057749 & 0.467589 \\ 0.730331 & 0.488609 \\ 0.952011 & 0.004614 \end{pmatrix}$$

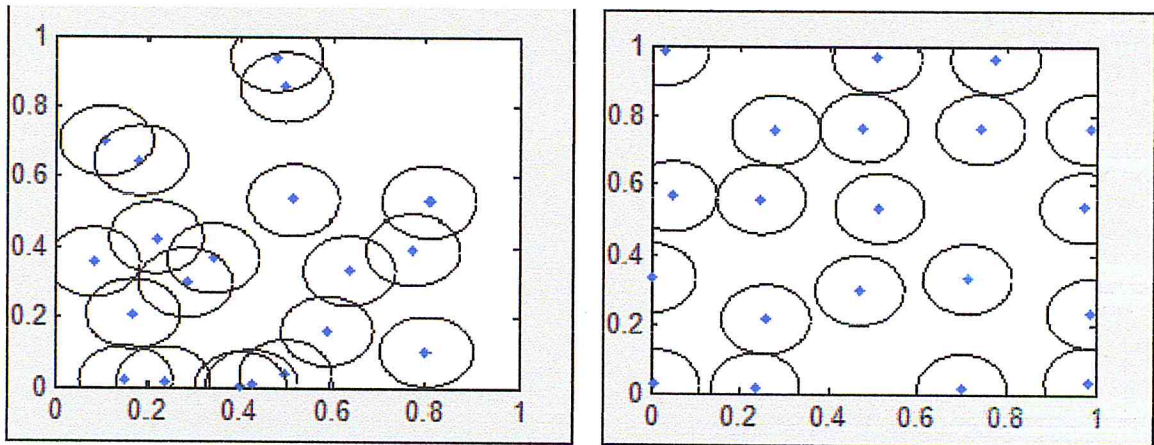


Figure (4.6): A gauche. Un plan initial. A droite. Un plan de Strauss obtenu
partir d'un potentiel puissance $\alpha = 0.5$

la figure de droite. Ainsi, l'utilisation d'un potentiel puissance permet de mieux répondre à la question du remplissage de l'espace.

Cependant, des alignements persistent comme on peut l'observer sur la figure de droite pour le facteur x_2 au niveau de la valeur 0.001, une solution est de combiner les contraintes sur les marges de dimension un avec le potentiel puissance afin d'obtenir une répartition uniforme sur les marges de dimension un et une répartition optimale au sens du remplissage de l'espace.

4.3 Etude des plans minimaux par Le rapport de distances:

1^{ere} cas:

$$n = 20 \quad N = 100$$

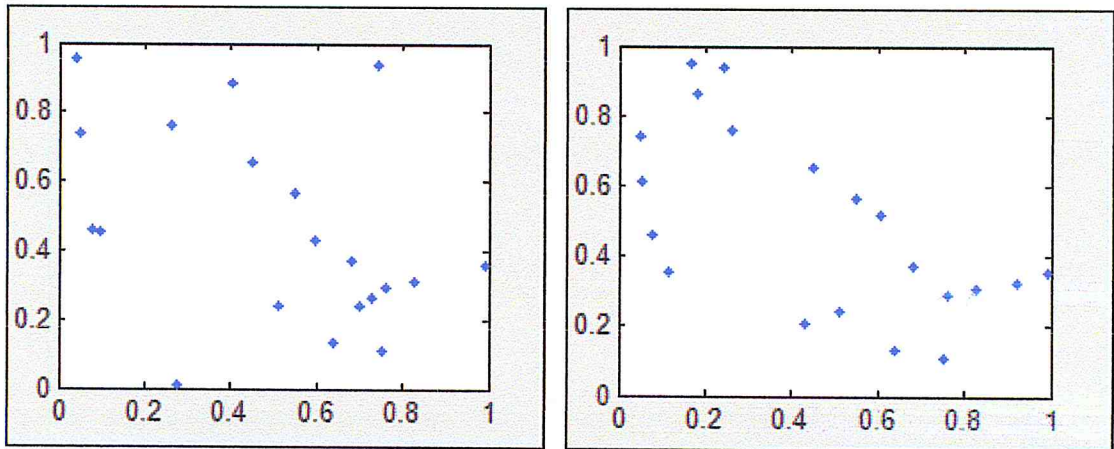


Figure (4.7) A gauche, un plan initiale pour $R = 19,72$ et A droite, un plan finale pour $R = 1,86$

2^{ere} cas:

$n = 20$ $N = 10000$

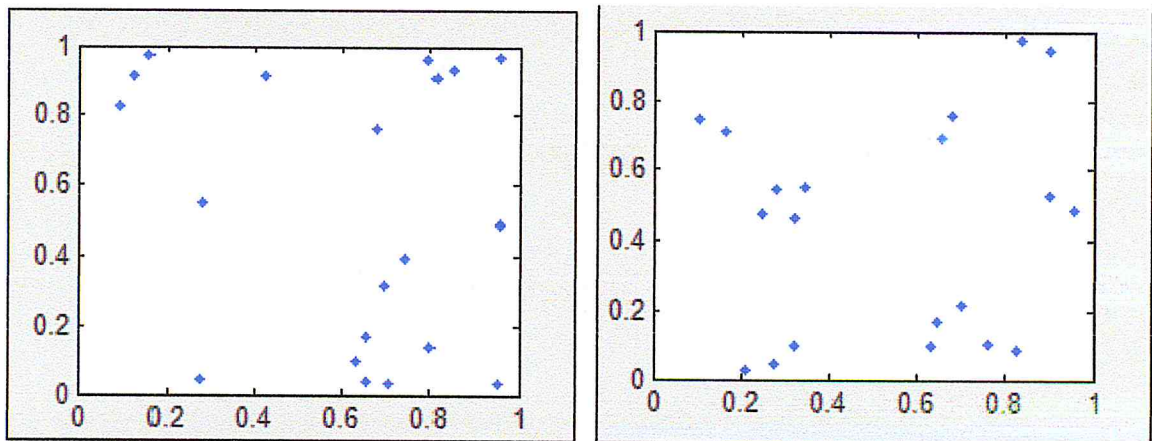


Figure (4.8) A gauche, un plan initiale pour $R = 18.2$ et A droite, un plan finale pour $R = 1.06$

Dans le plan initial le rapport de distance est plus grand et le plan final le rapport de distance est proche de 1, et dans ce cas les points sont proches de ceux d'une grille régulière.

4.4 Etude des plans à entropie maximale

$n = 20$ $MX = 10000$

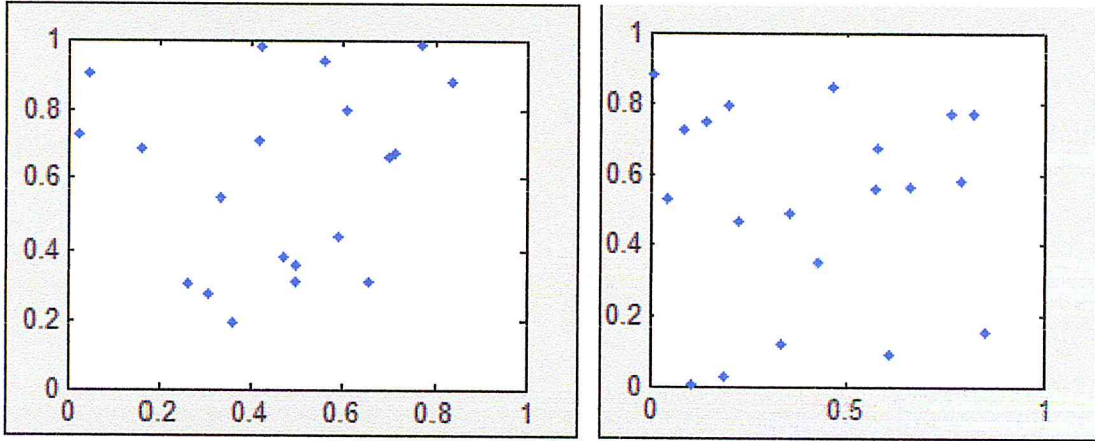


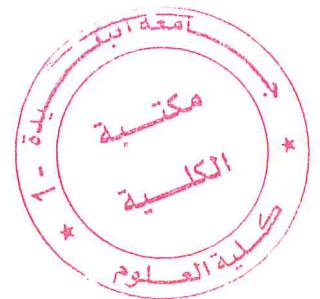
Figure (4.9). A gauche, un plan initial, alatoire, 20 points. A droite, un plan 20 points construit en maximisant l'entropie

Dans le plan initial l'entropie associée est minimale et la quantité d'information n'est pas disponible et dans le plan final l'entropie est maximale et la quantité d'information est disponible.

La Méthodologie de la Recherche Expérimentale, appelée aussi planification des expériences, est une discipline mathématique faisant partie de la statistique différentielle. La théorie de la méthode des plans d'expériences reste toujours en développement en utilisant des nouveaux critères de planification d'expériences. Cette thèse propose une nouvelle méthode de planification d'expériences basée sur les critères suivants : le critère de discrédance pour mesurer si la distribution donnée est proche d'une distribution uniforme, le critère de distance qui compare une distribution donnée à une Grille régulière et le critère d'entropie qui mesure la quantité d'information contenue dans le plan.

L'utilisation des méthodes MCMC dans le contexte des plans d'expériences numériques permet de construire de nouveaux plans spécifiés à partir d'une loi de Strauss et même plus généralement d'une loi de Gibbs. Cette approche offre une grande souplesse dans la mesure où l'on peut facilement jouer sur cette loi via sa représentation en termes de potentiels locaux d'énergie pour imposer des propriétés (par exemple de remplissage).

Le travail de ce mémoire a été par ailleurs de proposer et d'utiliser différents outils d'étude de la qualité des plans d'expériences pour s'assurer de leurs propriétés. Nous avons basé essentiellement nos recherches sur des outils de visualisation qui permettent de conclure rapidement sur la qualité des plans.



- [1] Philippe TRIBOULET (Lycée Niepce – Chalon sur Saône).NOTIONS DE BASES SUR LES PLANS D'EXPERIENCES 09/09/2008.
- [2] GOUPY Jacques. LES PLANS D'EXPERIENCES 2006.
- [3] Stéphane VIVIER.Stratégies d'optimition par la méthode des plans d'expériences et Application aux dispositifs électrotechniques modélisés par éléments finis
- [4] Jacques Goupy • Lee Creighton. INTRODUCTION AUX PLANS D'EXPÉRIENCES (3^e édition) 2006.
- [5] Hichem el Mossaoui. proposition de plans d'expériences dans la methodologie de la recherche experimentale:etude et programmation. Memoire de magister de mathematiques. Université de Blida. 2004.
- [6] Jean-Pierre Gauchi. Plans d'expériences optimaux : un expose didactique 2005
- [7] AZAZI Sabrina. Plans D-Optimaux Algorithmes d'échange et de déplacement 2016
- [8] jacques Goupy. (ingénieur-consiel) plans d'expéieces pour surfaces de réponde
- [9] Jessica FRANCO. Planification d'expériences numériques en phase exploratoire pour la simulation des phénomènes complexes 10/09/2008
- [10] Nils Berglund. Chaines de Markov. décembre 2007