

الجمهورية الجزائرية الديمقراطية الشعبية

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE SAAD DAHLAB - BLIDA

FACULTE DES SCIENCES

DEPARTEMENT DE MATHEMATIQUE



DOMAINE : MATHEMATIQUE ET INFORMATIQUE

FILIERE : MATHEMATIQUE

OPTION : RECHERCHE OPERATIONNELLE

Mémoire présenté pour l'obtention
Du diplôme de Master Académique



THEME

Méthodologie de la recherche Expérimentale

Plans numériques

Réalisé par :

- ❖ Benmesbah Yasmine
- ❖ Omar Abdelhak

proposé et encadré par :

ELMOSSAOUI HICHEM

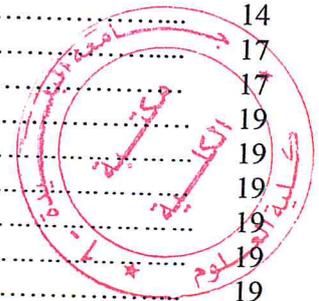
Président : Tamimi Omar
Examinateur : Dalila Haber

MA-510-62-1

Année universitaire : 2017 /2018

Table Des Matières

Résumé.....	vii
Table des matières.....	viii
Liste des symboles.....	ix
Liste des figures.....	xi
Liste des tableaux.....	viii
Liste des algorithmes.....	xiii
Introduction Générale.....	2
Chapitre 1 : Vers la théorie des plans d'expériences.....	4
1 La méthodologie de la recherche expérimentale.....	4
1.1 La méthode des plans d'expériences.....	5
1.2 Notion d'espace expérimental.....	5
1.2.1 Réponses, facteurs, niveaux.....	5
1.2.2 Domaine expérimental.....	7
1.2.3 Effet.....	8
1.2.4 Interactions.....	8
1.2.5 Surface de réponse.....	9
1.2.6 Coordonnées centrées réduites.....	10
1.2.7 Plans d'expériences.....	11
1.2.8 Matrice d'expériences.....	11
1.2.9 Performance et limites de la méthode.....	12
1.3 Notion de modélisation mathématique.....	14
1.4 Critères d'optimalité.....	17
1.4.1 Qualité de la représentation des coefficients.....	17
1.4.2 Critère A (<i>Average Variance of Parameters</i>).....	19
1.4.3 Critère D (<i>Déterminant</i>).....	19
1.4.4 Critère E (<i>Eigen Values</i>).....	19
1.4.5 Critère G (<i>General Variance</i>).....	19
1.4.6 <i>Efficacité-G</i>	19
1.4.7 Critère M (<i>critère du déterminant de la matrice des Moments</i>).....	19
1.4.8 Critère d'orthogonalité.....	20
1.4.9 Critère de presque orthogonalité.....	20
1.4.10 Critère d'iso variation par rotation.....	20
1.5 Plans d'expériences numériques usuels.....	21
1.5.1 Plans factoriels complets.....	21
1.5.2 Plans factoriels fractionnaires.....	21
1.5.3 Plans composites.....	22
1.5.4 Plans de Box-Behnken.....	23
1.5.5 Plans de Doehlert ou réseaux uniformes.....	23
1.6 Plans Marginaux.....	24
1.6.1 Hypercubes latins.....	24
1.6.2 Tableaux orthogonaux.....	25
1.7 Les Suites à faible discrédance.....	26



1.7.1 Suites de Halton	28
1.7.2 Suites de Hammersley.....	29
1.7.3 Suites de Sobol.....	30
1.7.4 Suites de Faure	31
1.7.5 Good Lattice Points	32
1.8 Les plans des mélanges	32
1.9 Récapitulatif.....	33
1.10 Présentation des méthodes d'évaluation de la qualité d'une structure.....	34
1.10.1 Critère Mindist.....	35
1.10.2 Critère Moydist.....	35
1.10.3 Critère de recouvrement.....	36
1.10.4 Critère de rapport de distances	37
1.10.5 Critère discrédance.....	37
Chapitre 2 : Simulation d'un processus ponctuel marqué par la méthode de Monte Carlo par Chaîne de Markov (M.C.M.C).....	39
2.1 Processus ponctuels.....	39
2.2 Processus ponctuels marqués.....	42
2.2.1 Quelques modèles.....	42
2.2.2 Processus de Markov	43
2.3 Simulation des processus ponctuels et convergence.....	44
2.4 Les algorithmes de simulation.....	47
2.4.1 L'algorithme de type Metropolis-Hastings (MH)	48
2.4.2 L'Algorithme de Metropolis-Hastings-Green (MHG).....	49
2.4.3 L'échantillonneur Gibbs	49
Chapitre 3 : Génération de nouveau plan à partir des processus ponctuels marqués.....	51
3.1 Plan marqué de Gibbs.....	52
3.2 Le choix des marques.	52
3.2.1 Fonction de variance de prédiction	53
3.2.2 Nombre de points marqués dans la configuration	53
3.3 Simulation des processus ponctuels par la méthode MCMC et l'algorithme de Metropolis- Hasting.....	55
3.4 Etude théorique de la planification d'expériences par méthodes McMC.....	53
Chapitre 4 : IMPLÉMENTATION, DISCUSSIONS ET COMPARAISON.....	60
4.1 Matlab comme outil de programmation.....	60
4.2 Etude de comparaison avec les plans usuels.....	61
4.3 Résultats et discussions.....	62
4.3.1 Résultat pour 10 points et 2 dimensions.....	64
4.3.2 Résultat pour 10 points et 3 dimensions.....	65
4.3.3 Résultat pour 50 points et 7 dimensions.....	65
Conclusion Générale.....	68
Remercîment	69
Bibliographie.....	70

Liste Des Symboles

a_0 :	Constante d'un modèle.
a_i :	Coefficients des termes de premier degré.
a_i^2 :	Coefficients des termes de deuxième degré.
a_{ij} :	Coefficients des termes carrés.
a_{ij} :	Coefficients des termes rectangles.
\hat{a}_i :	Coefficients de la variable x dans un modèle déterminé par la méthode des moindres carrés.
A :	Vecteur des coefficients du modèle.
\hat{A} :	Vecteur des coefficients estimés déterminés par la méthode des moindres carrés.
α :	Le niveau de confiance choisi.
$d(\hat{y}_u)$:	Fonction d'erreur de prédiction.
$d^2(\hat{y}_u)$:	Fonction de la variance de prédiction.
ddl :	Degré de liberté.
Δ :	L'écart d'ajustement ou « Lack of fit ».
e_i :	Ecart résiduel.
e :	Vecteur des écarts.
Eff-G :	Efficacité vis-à-vis du critère G.
f :	La fonction objective.
F :	Valeur de Fisher-Snedecor calculée.
I :	Matrice unité.
d :	Nombre de facteurs.
n :	désigne le nombre des expériences (les points).
N :	La loi Normale.
p :	Nombre de coefficients d'un modèle mathématique.
s^2 :	Estimation de la variance expérimentale.
X :	Matrice de calcul des effets.
X^T :	Matrice transposée de X.
$(X^T X)$:	Matrice d'information.
$(X^T X)^{-1}$:	Matrice de dispersion.
y_i :	Réponse de l'essai n°i.
y :	Réponse prédite de l'essai i.
\hat{y} :	Vecteur des réponses prédites.
t :	désigne la force d'un tableau orthogonal.
\hat{F}_n :	désigne la fonction de répartition empirique de n expériences.
F_u :	désigne la fonction de répartition de la loi uniforme de n points sur [0,1] .

(x_n)	désigne Une suite <i>aléatoire</i> distribuée sur le domaine d'expérience.
$D_n(x)$	désigne la discrétance qui permet de mesurer l'irrégularité de la distribution de (x_n) .
$V(f)$	désigne la variation d'une fonction f .
$\phi_b(i)$	désigne La fonction radicale inverse en base b .
C_p	désigne la suite de Van Der Corput en base b .
$\text{dist}(x^i, x^k)$	désigne la distance euclidienne entre deux points x^i et x^k .
λ	désigne la mesure de recouvrement de l'espace par les points.
(χ, d)	un espace métrique, χ définissant l'espace muni d'une métrique d
x	Une configuration non ordonné de n points en χ .
$\aleph_x(A)$	Le nombre fini de points de x dans A .
(Ω, A, \mathcal{P})	un espace probabilisé.
X	La processus ponctuel
\aleph_x	la famille des configurations de points de χ localement finies
$\pi(\cdot)$	La loi d'un processus conditionnellement au nombre de points donnée.
$S(x)$	est le nombre d'interaction entre chaque paire de point.
k	la constante de normalisation.
γ	un coefficient de répulsion.
$U(x)$	est la fonction de potentielle.
$a(x)$	est l'intensité du processus i.e. le nombre moyen d'objets de x .
$C(x)$	Le nombre de composante connexe dans x .
P	Défini le noyau de transition.
$C(K)$	Le coefficient de contraction (coefficient Dobrushin 1956).
$R(x, y)$	le rapport d'acceptation dans l'algorithme de type Metropolis-Hastings (MH)
β	un paramètre d'échelle, qui corrige l'intensité du processus

Liste Des Figures

Figure 1.1	Domaine de variation du facteur.....	5
Figure 1.2	Définition de l'espace expérimental.....	6
Figure 1.3	Niveaux des facteurs définissant des points expérimentaux dans l'espace expérimental.....	6
Figure 1.4	Définition d'un domaine d'étude.....	7
Figure 1.5	Effet d'un facteur.....	8
Figure 1.6	Interactions des facteurs x_1 et x_2 dans trois cas.....	9
Figure 1.7	Définition de la surface de réponse.....	10
Figure 1.8	La théorie des plans d'expériences montre que les meilleurs emplacements sont les sommets A,B,C et D du domaine d'étude.....	11
Figure 1.9	Ellipsoïde de confiance.....	18
Figure 1.10	Plans composites pour l'étude de deux facteurs.....	21
Figure 1.11	Plans composites pour trois facteurs.....	22
Figure 1.12	Plans Box-Behnken pour trois facteurs ...	23
Figure 1.13	Plans Dohlert pour deux et trois facteurs.....	24
Figure 1.14	Un hyper cube latin à 5 points en 2 dimensions.....	25
Figure 1.15	montre les points générés par un tableau orthogonal de type OA_1 projetés sur le sous-espace (X_1, X_2)	26
Figure 1.16	De la gauche vers la droite les 50, 250 et 500 premiers points d'une suite de Halton en base 2 et 3.....	29
Figure 1.17	De la gauche vers la droite les 50, 250 et 500 premiers points d'une suite de Hammersley en base 2.....	30
Figure 1.18	les 50, 250 et 500 premiers points d'une suite de Sobol en dimension.....	31
Figure 1.19	les 50, 250 et 500 premiers points d'une suite de Faure en dimension 2.....	32
Figure 1.20	Comparaison de trois distributions de points. a) distribution uniforme avec deux points très proches, b) tous les points sont proches et constituent un seul amas, c) distribution uniforme.....	36
Figure 2.1	une configuration de points dans $[0, 1]^2$	40
Figure 2.2	Plan de Strauss à 20 points en dimension 2 et $\gamma = 0.2$ et de rayon=0.01... 41	41
Figure 2.3	plan de Gibbs à 20 points en dimension 2 et de rayon = 0.01.....	41
Figure 2.4	Une réalisation de l'ensemble $B(x)$ permettant de définir l'aire $a(x)$ et le nombre de composantes connexes $c(x)$	44
Figure 2.5	les familles d'algorithme pour la simulation des chaînes MCMC.....	47
Figure 3.1	La fonction de potentielle puissance pour différents α et $R=0.2$	48
Figure 4.1	Box plots des critères de qualité pour deux dimensions.....	63
Figure 4.2	Box plots des critères de qualité pour trois dimensions.....	65
Figure 4.3	Box plots des critères de qualité pour sept dimensions.....	67

Liste Des Algorithmes

algorithme 2.1	Metropolis Hastings	48
algorithme 2.2	Metropolis Hastings Green	49
algorithme 2.3	Echantillonneur de Gibbs	50
algorithme 3.1	Métropolies Hastings Green, simulation de processus ponctuel marqué de Gibbs.....	55

ملخص

إن طريقة التصميم التجريبي هي مجموعة من أساليب وطرق التفكير المنطقي لأي مجرب يرغب في إجراء تخطيط تجريبي ، وفي هذا السياق ، نحن مهتمون ببناء تصميمات تجريبية من محاكيات حتمية وفي المرحلة الاستطلاعية أي عندما يكون الاعتماد بين المدخلات والمخرجات غير معروف سلفاً.

في هذا العمل نقترح طريقة جديدة لتخطيط التجربة تعتمد على محاكاة احتمالية للعينات بواسطة طريقة مونت كارلو

سلسلة ماركوف (MCMC).

وقد تم النظر في عمليات نقاط غيبس المعلمة (مع تفاعلات نوع التنافر) من أجل تلبية المتطلبات المطلوبة (ملء الفراغات

وتوزيع النقاط بشكل موحد).

RESUME

La méthode des plans d'expériences est un ensemble de méthode et de mode de raisonnement destiné à tout expérimentateur désirant faire de la planification expérimentale, dans ce contexte nous nous sommes intéressés à la construction des plans d'expériences à partir de simulateurs déterministes et en phase exploratoire *i.e.* lorsque la dépendance entre les entrées et les sorties est à priori inconnue.

Dans ce travail nous proposons 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). Les processus **ponctuels marqués** de Gibbs (avec interactions de type répulsion) ont été envisagés afin de répondre aux exigences souhaitées (le remplissage de l'espace et la distribution uniformité des points).

INTRODUCTION GENERALE

La simulation numérique est devenue un outil incontournable en recherche et développement pour évaluer par exemple des modèles d'expansion de l'univers, ou pour approcher des solutions optimales d'un stockage de déchets radioactifs ou d'une architecture d'un barrage. Les codes sont constitués de modèles numériques reliant des variables, descriptives de l'état du système simulé, à un certain nombre de paramètres caractérisant, par exemple, les conditions initiales du système ou bien certaines lois ou relations utilisées dans la modélisation. Le simulateur prend ces paramètres en entrée et délivre en sortie les réponses sur les variables d'état du système. Une simulation consiste à fixer un ou plusieurs jeux de variables d'entrée, à réaliser les calculs, puis à analyser les réponses fournies par le simulateur. Malgré les progrès toujours croissants de l'informatique, ces codes de calcul sont complexes et souvent très longs à exécuter. L'utilisateur souhaite, dans ce cas, obtenir un modèle plus simple, plus rapide et qui rende compte le mieux possible des réponses du simulateur.

La construction du modèle simplifié est fondée sur la réalisation d'un certain nombre de simulations appelées **expériences numériques**. Les expériences simulées étant coûteuses en temps de calcul, il est important de disposer de méthodes permettant d'optimiser leur planification. Pour une famille de fonctions retenues pour construire le modèle simplifié, la qualité de représentation du modèle simplifié dépendra principalement du choix de ces expériences.

Des méthodes de planification d'expériences ont été étudiées par des auteurs comme Fisher [1962] [1972] ou Box [1960]. Leurs travaux ont donné lieu à une nouvelle branche des statistiques connue sous le nom de méthodologie des plans d'expériences. La plupart des critères de ces plans utilisent le bruit de mesure qui existe dans le cadre d'expériences réelles. Les expériences simulées ne rentrent pas dans ce contexte de planification d'expériences car les simulations que l'on considère pour créer le modèle simplifié sont déterministes. Si l'on effectue plusieurs simulations avec des jeux de variables d'entrée identiques, on obtient la même réponse. On ne peut donc pas utiliser directement la méthodologie des plans d'expériences. On peut toutefois l'adapter en considérant que ce bruit expérimental tend vers zéro.

CHAPITRE 1

VERS LA THEORIE DES PLANS D'EXPERIENCES

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 et 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 Terminologie des plans d'expériences

1.1 La méthodologie de la recherche expérimentale

La méthodologie de la Recherche Expérimentale est utile à toutes les personnes qui entreprennent des recherches scientifiques où des études industrielles. L'emploi des plans d'expériences pour l'étude empirique d'une loi de réponse pose aux statisticiens ou aux chercheurs des problèmes particuliers. Alors qu'ils ont peu d'informations sur cette loi, ils ne peuvent disposer en général que d'un échantillon d'observations très limité au regard du nombre des paramètres des modèles qu'ils peuvent envisager pour leurs analyses. Avant toute observation de la réponse il doit donc préciser, non seulement quels modèles utiliser, mais encore, comment organiser les expériences. En effet, la qualité de l'analyse statistique dépend étroitement du plan expérimental utilisé pour observer la réponse à étudier. Par ailleurs, il faut généralement recourir à l'analyse combinatoire pour construire les plans d'expériences proposés.

Pour proposer une solution répondant aux objectifs industriels, il est donc 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. Il est

CHAPITRE 1

VERS LA THEORIE DES PLANS D'EXPERIENCES

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 et 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 Terminologie des plans d'expériences

1.1 La méthodologie de la recherche expérimentale

La méthodologie de la Recherche Expérimentale est utile à toutes les personnes qui entreprennent des recherches scientifiques où des études industrielles. L'emploi des plans d'expériences pour l'étude empirique d'une loi de réponse pose aux statisticiens ou aux chercheurs des problèmes particuliers. Alors qu'ils ont peu d'informations sur cette loi, ils ne peuvent disposer en général que d'un échantillon d'observations très limité au regard du nombre des paramètres des modèles qu'ils peuvent envisager pour leurs analyses. Avant toute observation de la réponse il doit donc préciser, non seulement quels modèles utiliser, mais encore, comment organiser les expériences. En effet, la qualité de l'analyse statistique dépend étroitement du plan expérimental utilisé pour observer la réponse à étudier. Par ailleurs, il faut généralement recourir à l'analyse combinatoire pour construire les plans d'expériences proposés.

Pour proposer une solution répondant aux objectifs industriels, il est donc 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. Il est

nécessaire d'utiliser une approche méthodologique qui permet non seulement de réduire le coût de l'expérimentation, mais aussi d'établir une organisation optimale des expériences.

La méthodologie "désigne l'ensemble des méthodes et outils qui peuvent être utilisés pour poser le problème, diriger l'expérimentation et en exploiter les résultats. Comme nous le verrons plus loin, elle comporte aussi la notion de "Stratégie expérimentale". Le terme "Expérimentale" signifie que l'expérience est la seule voie d'obtention des informations non disponibles ou inexistantes.

1.2 La méthode des plans d'expériences

Les plans d'expériences permettent l'étude de nombreux facteurs tout en maintenant le nombre des essais à des valeurs raisonnables. Une de leurs applications principales est la recherche des facteurs influents. La compréhension de la méthode des plans d'expériences s'appuie sur deux notions essentielles, celle d'espace expérimental et celle de modélisation mathématique des grandeurs étudiées [18].

1.2.1 Notion d'espace expérimental

1.2.1.1 Réponses, facteurs, niveaux

- 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 1.1). La valeur donnée à un facteur pour réaliser un essai est appelée niveau. Lorsque l'on étudie l'influence d'un facteur, en général on limite ses variations entre deux bornes (borne inférieure c'est le niveau bas ; borne supérieure c'est le niveau haut).

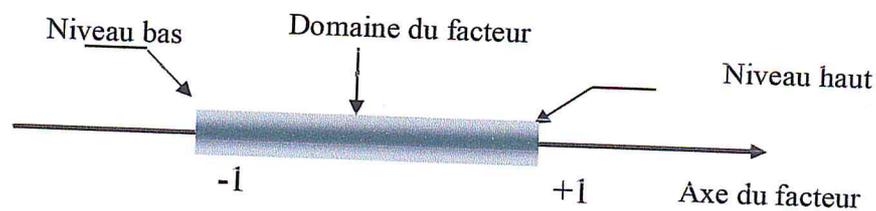


Figure 1.1 : Domaine de variation du facteur

- L'ensemble de toutes les valeurs que peut prendre le facteur entre le niveau bas et le niveau haut, s'appelle le domaine de variation du facteur. On a l'habitude de noter le niveau bas par -1 et le niveau haut par $+1$.
- S'il y a un second facteur, il est représenté, lui aussi, par un axe. On définit, comme pour le premier facteur, son niveau bas, son niveau haut et son domaine de variation. Ce second axe est disposé orthogonalement au premier. On obtient ainsi un repère cartésien qui définit un espace euclidien à deux dimensions. Cet espace est appelé l'espace expérimental (Figure 1.2).

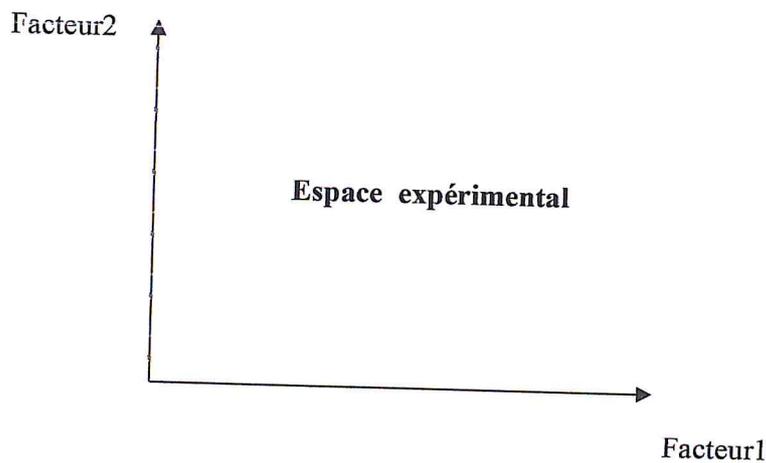


Figure 1.2 : Définition de l'espace expérimental

- Le niveau x_1 du facteur 1 et le niveau x_2 du facteur 2 peuvent être considérés comme les coordonnées d'un point de l'espace expérimental (Figure 1.3).

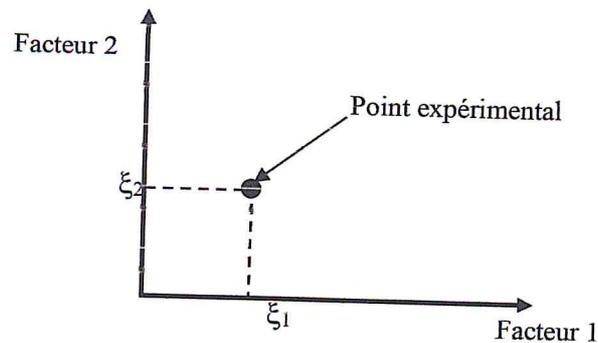


Figure 1.3 : Niveaux des facteurs définissant des points expérimentaux dans l'espace expérimental

Une expérience donnée est alors représentée par un point dans ce système d'axes. Un plan d'expériences est représenté par un ensemble de points expérimentaux.

1.2.1.2 Domaine expérimental

Nous définissons le domaine de validité de l'expérience comme étant les limites raisonnables de variation des facteurs. En effet, peu importe le comportement des facteurs à l'extérieur de ce domaine, ce qui nous intéresse c'est le comportement des facteurs à l'intérieur.

Nous allons choisir les niveaux des facteurs pour couvrir le domaine de validité des facteurs. On choisit les limites de ce domaine de telle sorte qu'il soit le plus large possible, tout en s'assurant que le fonctionnement du système sera relativement homogène dans ce domaine.

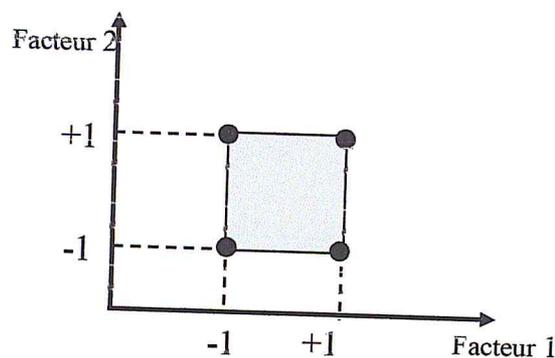


Figure 1.4 : Définition du domaine d'étude

Les deux niveaux choisis sont l'objet de compromis entre deux risques :

- Si les niveaux sont trop proches l'un de l'autre, on risque de ne pas mettre en évidence l'effet des facteurs.
- Si les niveaux sont trop éloignés, l'hypothèse de linéarité est moins réaliste, et on risque, en plus, de tomber sur des combinaisons irréalisables.

Le domaine d'étude (ou domaine expérimental) est la partie de l'espace expérimental retenue par l'expérimentateur pour faire ses essais. Une étude, c'est-à-dire un ensemble d'expériences bien définies, est représentée par une série de points

disposés dans le domaine d'étude. Cette manière de représenter une expérimentation par un ensemble de points dans un espace cartésien est une représentation géométrique de l'étude.

1.2.1.3 Effet

L'effet global d'un facteur x correspond à la variation de la réponse y lorsque x passe d'une valeur -1 à une autre valeur $+1$ comme présenté sur la figure 1.5. Graphiquement, plus l'inclinaison est forte, plus l'effet est important, ce qui peut déjà donner des indications. On appelle effet moyen d'un facteur (sous-entendu : sur la réponse) la demi-variation de la réponse quand le facteur passe du niveau -1 au niveau $+1$. Ainsi, l'effet moyen est défini comme étant la moitié de l'effet global.

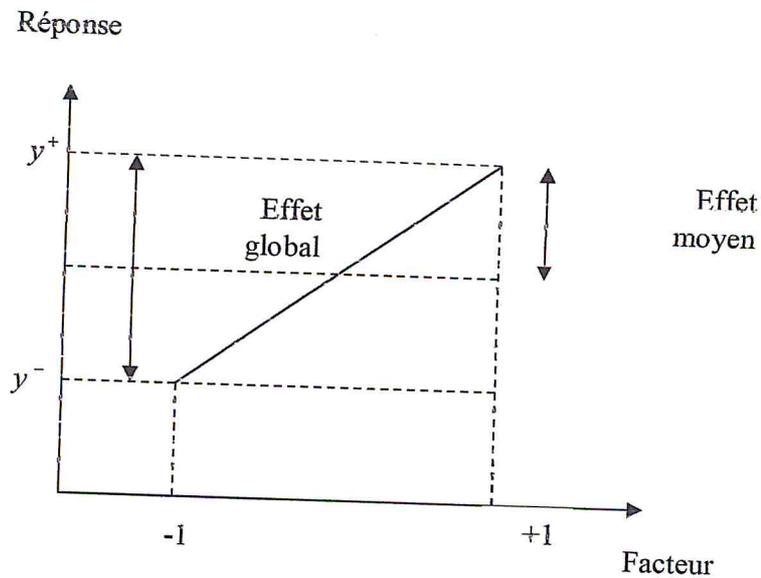


Figure 1.5 : Effet d'un facteur.

1.2.1.4 Interactions

Deux facteurs x_1 et x_2 sont en interaction si les effets de l'un dépendent du niveau de l'autre et réciproquement. Elle est notée $x_1 x_2$, et son nombre de niveaux est le produit des nombres de niveaux de x_1 et x_2 . La figure 1.6 montre graphiquement l'interaction de deux facteurs x_1 et x_2 , de gauche à droite, d'abord sans interaction, ensuite avec une faible interaction enfin avec une forte interaction.

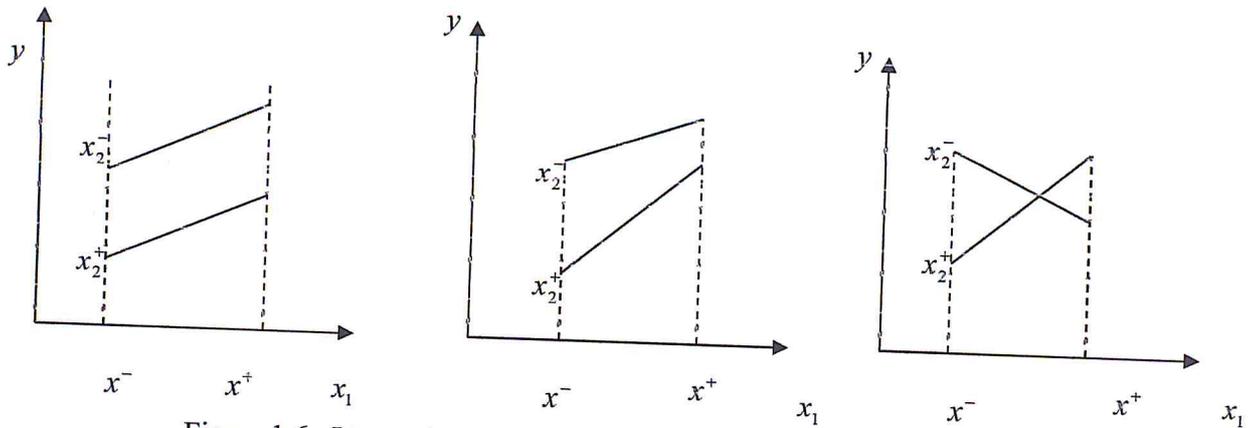


Figure 1.6 : Interactions des facteurs x_1 et x_2 dans trois cas

Nous appellerons «interactions d'ordre deux», les interactions intervenant entre deux facteurs. De même, nous appellerons interactions d'ordre trois, celles intervenant entre trois facteurs.

1.2.1.5 Surface de réponse

Les niveaux x_i représentent les coordonnées d'un point expérimental et y est la valeur de la réponse en ce point. On définit un axe orthogonal à l'espace expérimental et on l'attribue à la réponse. La représentation géométrique du plan d'expériences et de la réponse nécessite un espace ayant une dimension de plus que l'espace expérimental. Un plan à deux facteurs utilise un espace à trois dimensions pour être représenté : une dimension pour la réponse, deux dimensions pour les facteurs.

A chaque point du domaine d'étude correspond une réponse. A l'ensemble de tous les points du domaine d'étude correspond un ensemble de réponses qui se localisent sur une surface appelée la surface de réponse (Figure 1.7). Le nombre et de l'emplacement des points d'expériences est le problème fondamental des plan d'expériences. On cherche à obtenir la meilleure précision possible sur la surface de réponse tout en limitant le nombre d'expériences.

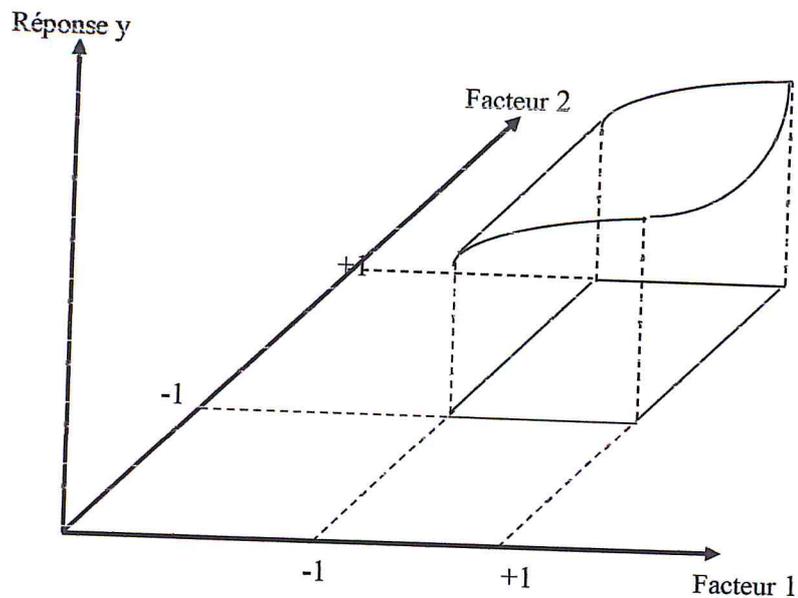


Figure 1.7 : Définition de la surface de réponse

1.2.1.6 Coordonnées 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
- on change l'unité des mesures

Ces deux modifications entraînent l'introduction de nouvelles variables que l'on appelle variables centrées réduites, centrées pour indiquer le changement d'origine et réduites pour signaler la nouvelle unité. Le passage des variables d'origine z aux variables centrées réduites x (variables sans dimensions), et inversement, est donné par la formule suivante :

$$x = \frac{z - z_0}{pas}$$

Où $z_0 = (\text{niveau haut} + \text{niveau bas})/2$

$pas = (\text{niveau haut} - \text{niveau bas})/2$

L'intérêt des variables centrées réduites est de pouvoir présenter les plans d'expériences de la même manière quelque soient les domaines d'études retenus et

quelque soient les unités des facteurs. La théorie des plans d'expériences présente ainsi une grande généralité.

1.2.1.7 Plans d'expériences

Chaque point du domaine d'étude représente des conditions opératoires possibles, donc une expérience que l'opérateur peut réaliser.

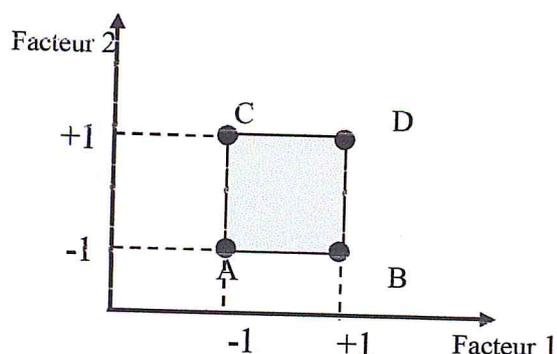


Figure 1.8 : La théorie des plans d'expériences montre que les meilleurs emplacements sont les sommets A, B, C et D du domaine d'étude

Le choix du nombre et de l'emplacement des points d'expériences est le problème fondamental des plans d'expériences. On a l'habitude d'appeler plans d'expériences des ensembles de points expérimentaux répondant à des propriétés bien précises. Ce sont les plans d'expériences classiques. Ils sont connus et largement publiés. Lorsque les points expérimentaux sont disposés autrement que dans les plans d'expériences classiques, on parle de plans non conventionnels. Leurs propriétés sont le plus souvent moins bonnes que celles des plans classiques. Mais ce sont des plans que l'on rencontre car il n'est pas toujours possible de respecter les impératifs des plans d'expériences classiques.

1.2.1.8 Matrice d'expériences

La représentation géométrique d'un plan d'expériences est commode pour imaginer la position des points expérimentaux dans le domaine d'étude. Mais elle ne peut plus être employée dès que le nombre de facteurs est supérieur à deux. Pour les espaces multidimensionnels, nous adopterons une représentation sous forme de tableau ou matrice d'expériences. La matrice d'expériences est un objet mathématique qui représente, sous forme codée ou normée, l'ensemble des expériences à réaliser.

C'est un tableau constitué de N lignes correspondant aux N expériences à réaliser et de k colonnes, correspondant aux k variables étudiées. L'élément x_{ij} de la matrice ainsi formée correspond à la valeur du niveau que prend la $j^{\text{ième}}$ variable à la $i^{\text{ième}}$ expérience. La matrice d'expériences (Tableau 1.1) définit les essais à réaliser. Le terme essai est l'équivalent de point d'expérience.

Tableau 1.1 : Matrice d'expériences

N° essai	Facteur 1	Facteur 2
1(A)	-1	-1
2(B)	+1	-1
3(C)	-1	+1
4(D)	+1	+1

1.3 Performance et limites de la méthode

La méthode traditionnelle d'expérimentation qui consiste à faire varier les facteurs étudiés l'un après l'autre, présente les failles suivantes :

- Les essais sont réalisés d'une manière séquentielle, sans planification préalable de l'ensemble, et l'interprétation des résultats se fait pas à pas cela conduit à une campagne expérimentale d'une durée longue et aléatoire, et jette le doute sur la solution retenue.
- Les facteurs sont variés l'un après l'autre, un facteur à la fois, ce qui ne permet pas la mise en évidence des interactions (couplage des effets) entre ces facteurs.
- La difficulté de la mise en évidence de la variabilité naturelle des grandeurs étudiées (réponses) occasionnée par les différentes sources de fluctuation.

Cela entache les résultats, les prévisions et les conclusions d'une incertitude qui peut être importante. Dans la méthode des plans d'expériences, la réflexion profonde dans la planification des essais, permet de réfléchir moins et de se poser moins de questions après la réalisation de ces essais.

La méthode puise son efficacité dans l'utilisation de deux outils complémentaires :

1. L'outil algébrique permet de faire varier tous les facteurs en même temps (à la fois) pour tenir compte de leur couplage. Cela permet de gagner de la précision dans les résultats ; la modélisation et les prévisions.
2. L'outil statistique permet de faire la plupart de ce qui est du à la variabilité Naturelle ; et de ce qui est du aux effets réels des facteurs, car les résultats d'essais ne sont jamais parfaitement reproductible (grandeurs aléatoires).

Cette méthode permet de répondre aux questions suivantes :

- Parmi les facteurs retenus à priori, quels sont ceux qui ont une influence réelle sur la réponse étudiée ? (**Criblage**)
- Parmi les facteurs ayant une influence réelle, quels sont ceux qui sont réellement importants ? (**Hierarchisation**)
- Quelles valeurs (nominale ou tolérance) donner aux facteurs influent pour que la réponse soit optimisée ? (**Optimisation**)
- Comment minimiser, voir neutraliser les effets indéfrisables de facteurs non maîtrisables ou coûteux à maîtriser (facteurs bruits) ? (**Robustesse**)

Cependant, plusieurs contraintes peuvent restreindre son application, à savoir :

- Les résultats des essais doivent pouvoir être quantifiés.
- Les facteurs doivent pouvoir être maîtrisés lors de l'expérimentation.
- Le système étudié doit être statistiquement stable, c'est-à-dire avoir une variabilité naturelle constante dans le temps.
- Les perturbations des procédures doivent être acceptables.

1.4 Notion de modélisation mathématique

Le modèle est une relation entre les facteurs x_1, x_2, \dots, x_k et la réponse que l'on désire étudier. La forme générale de modèle est représentée par l'expression suivante :

$$y = f(x_1, x_2, x_3, x_4, \dots, x_k).$$

Mais ce modèle pose deux questions. Le premier est : « Ce modèle, choisi avant de commencer les essais, est-il vraiment celui qui représente le mieux le phénomène étudié ? ». En réalité, il y a peu de chances de répondre par l'affirmative. Le modèle choisi a priori a peu de chances d'être le bon. Il faudra donc tenir compte de ce décalage si l'on veut que le modèle choisi se rapproche du modèle profond qui représente la réalité expérimentale et que nous ne connaissons pas. La seconde question est : « Comment tenir compte de l'erreur expérimentale dans le modèle ? ». On répond à la première question en ajoutant un terme d'ajustement au modèle initial et à la deuxième question en ajoutant un terme présentant l'erreur expérimentale. Le modèle prend la forme générale :

$$y = f(x_1, x_2, x_3, x_4, \dots, x_k) + \Delta + \varepsilon$$

Où :

- $f(x_1, x_2, x_3, x_4, \dots, x_k)$ est la partie mathématique du modèle. C'est le modèle que nous choisissons ;
- Δ est le terme d'ajustement (LOF : Lack of fit) qui représente la différence entre le modèle choisi et le modèle qui représente la réalité expérimentale ;
- ε est l'erreur expérimentale (pure error).

On a l'habitude de regrouper le terme d'ajustement et l'erreur expérimentale dans un seul terme que l'on appelle l'écart (error) et que l'on note par la lettre e .

Le modèle s'écrit donc :

$$y = f(x_1, x_2, x_3, x_4, \dots, x_k) + e$$

Dans le cadre des plans d'expériences, on utilise des modèles polynomiaux. Pour rester concret, nous allons reprendre le modèle polynomial du chapitre précédent en lui adjoignant le terme d'écart, soit :

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_{11}x_1^2 + a_{22}x_2^2 + e$$

Nous allons reprendre le raisonnement que nous avons suivi au chapitre précédent pour poser le système d'équation qu'il faut résoudre à l'issue d'un plan d'expériences.

Un premier essai effectué au niveau $x_{1,1}$ du facteur 1 et au niveau $x_{2,1}$ du facteur 2, apporte une première équation :

$$y_1 = a_0 + a_1x_{1,1} + a_2x_{2,1} + a_{12}x_{1,1}x_{2,1} + a_{11}x_{1,1}^2 + a_{22}x_{2,1}^2 + e_1$$

Un second essai effectué au niveau $x_{1,2}$ du facteur 1 et au niveau $x_{2,2}$ du facteur 2, apporte une seconde équation :

$$y_2 = a_0 + a_1x_{1,2} + a_2x_{2,2} + a_{12}x_{1,2}x_{2,2} + a_{11}x_{1,2}^2 + a_{22}x_{2,2}^2 + e_2$$

Le dernier essai effectué au niveau $x_{1,n}$ du facteur 1 et au niveau $x_{2,n}$ du facteur 2, apporte une dernière équation :

$$y_n = a_0 + a_1x_{1,n} + a_2x_{2,n} + a_{12}x_{1,n}x_{2,n} + a_{11}x_{1,n}^2 + a_{22}x_{2,n}^2 + e_n$$

A la fin du plan d'expériences, on obtient le système suivant :

$$\begin{cases} y_1 = a_0 + a_1x_{1,1} + a_2x_{2,1} + a_{12}x_{1,1}x_{2,1} + a_{11}x_{1,1}^2 + a_{22}x_{2,1}^2 + e_1 \\ y_2 = a_0 + a_1x_{1,2} + a_2x_{2,2} + a_{12}x_{1,2}x_{2,2} + a_{11}x_{1,2}^2 + a_{22}x_{2,2}^2 + e_2 \\ y_3 = a_0 + a_1x_{1,3} + a_2x_{2,3} + a_{12}x_{1,3}x_{2,3} + a_{11}x_{1,3}^2 + a_{22}x_{2,3}^2 + e_3 \\ \dots \\ y_n = a_0 + a_1x_{1,n} + a_2x_{2,n} + a_{12}x_{1,n}x_{2,n} + a_{11}x_{1,n}^2 + a_{22}x_{2,n}^2 + e_n \end{cases}$$

Dans cet ensemble de relations sont connues :

- Les différentes réponses y_i ,
- Les différentes coordonnées des points expérimentaux, $x_{1,i}, x_{2,j}$ ne sont pas connus,
- Les coefficients $a_0, a_1, a_2, a_{12}, a_{11}$ et a_{22} ,
- Les valeurs de n écarts e_i , encore appelés résidus.

Nous avons déjà rencontré le même type de problème et nous connaissons la plupart des matrices qui vont intervenir :

- La matrice des réponses (matrice de dimensions $n \times 1$)

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_i \\ \dots \\ y_n \end{bmatrix}$$

1.5 Critères d'optimalité

Les erreurs sur les réponses prédites sont liées à la matrice X . Comme cette matrice est connue, il est possible d'évaluer la qualité de l'expérimentation avant le début des expériences. La matrice X dépend du modèle mathématique postulé et de l'emplacement des points expérimentaux. Un bon plan d'expériences sera un plan permettant d'obtenir la plus faible erreur sur les réponses prédites. La règle que l'on peut suivre est d'obtenir une erreur sur les réponses prédites de l'ordre de grandeur de l'erreur sur les réponses mesurées. Selon les critères choisis, l'emplacement des points expérimentaux pourra varier d'un plan à l'autre. Il y a plusieurs critères d'optimalité. Il y a ceux qui s'intéressent à la répartition de la variance dans le domaine d'étude. C'est par exemple le critère d'iso variance par rotation. Il y a les critères d'optimalité qui ont pour objectif d'obtenir un modèle mathématique de bonne qualité. Ces derniers critères s'intéressent à la précision des coefficients du modèle. [8]

1.5.1 Qualité de la représentation des coefficients

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

$$(A - \hat{A})^T (X^T X) (A - \hat{A}) = ps^2 F_\alpha(p, ddl)$$

Où : A est le vecteur des coefficients du modèle,

\hat{A} est le vecteur des estimateurs des coefficients,

s^2 est une estimation de la variance expérimentale,

F est la statistique correspond au test de Fisher,

α est le niveau de confiance choisi,

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

p : est le nombre de coefficients du modèle,

Ce domaine de confiance 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 :

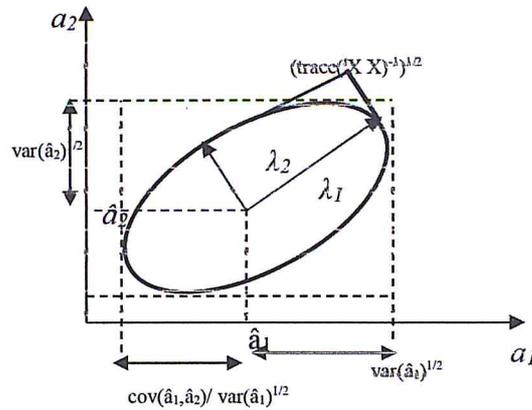


Figure 1.9 : Ellipsoïde de confiance

Où (\hat{a}_1, \hat{a}_2) sont les deux coefficients trouvés par la régression multilinéaire, (λ_1, λ_2) les deux valeurs propres de l'ellipsoïde [8]. Cet hyper ellipsoïde représente les performances de l'ensemble des observations pour un modèle donné. Il peut être décrit par 3 caractérisations :

- **Volume**

Le volume de l'ellipsoïde est égal à $\pi(\lambda_1 \lambda_2)^{1/2} = \pi(\det(X^T X)^{-1})^{1/2}$. Ce volume est lié au déterminant de la matrice de dispersion $(X^T X)^{-1}$. Plus le volume sera petit, plus l'hyper ellipsoïde tend vers un point et plus on se rapprochera de la solution vraie A .

- **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.

- **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 (2.7), l'hyper ellipsoïde dépend de la matrice de $(X^T X)$, donc les critères sont fondés sur les matrices d'information $(X^T X)$ et de dispersion $(X^T X)^{-1}$.

1.5.1.1 Critère A (Average Variance of Parameters)

Une matrice d'expériences est dite *A-optimale*, si la trace de $(X^T X)^{-1}$ est minimum. La précision sur l'ensemble des coefficients est améliorée.

1.5.1.2 Critère D (Déterminant)

Une matrice d'expériences est dite *D-optimale*, si elle conduit au déterminant minimal pour sa matrice de dispersion $(X^T X)^{-1}$.

1.5.1.3 Critère E (Eigen Values)

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

1.5.1.4 Critère G (General Variance)

Le critère-G (ou minimax) prend en compte la plus grande valeur, sur tout le domaine expérimental, de la fonction de variance $d^2(\hat{y}_p)$ engendré par la matrice X .

$$d_{\max} = \text{Max}(d^2(\hat{y}_p))$$

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

1.5.1.5 Efficacité-G

L'efficacité-G d'une matrice d'expériences est donnée par la formule suivante :

$$\text{Eff} - G = \frac{p}{n d_{\max}} 100$$

La matrice d'expériences est dite *optimale-G*, si parmi toutes les matrices à n expériences elle est celle qui conduit à la plus petite valeur du critère G.

1.5.1.6 Critère M (critère du déterminant de la matrice des Moments)

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 :

$$M = \frac{(X^T X)}{n}$$

Grâce à ce critère on peut comparer deux matrices d'expériences, n'ayant pas forcément le même nombre d'expériences. Soient M_1 et M_2 deux matrices des moments associées à deux matrices d'expériences constituées respectivement de N_1 et N_2 expériences.

$$M_1 = \frac{(X_1^T X_1)}{N_1}$$

$$M_2 = \frac{(X_2^T X_2)}{N_2}$$

Nous dirons que le premier plan est plus efficace que le deuxième plan vis-à-vis du critère M si :

$$> |M_2| |M_1|$$

1.5.1.7 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 $(X^T X)^{-1}$ (ou $X^T X$) est diagonale, donc quand les covariances des coefficients sont nulles.

1.5.1.8 Critère de presque orthogonalité

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

1.5.1.9 Critère d'iso variation 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 (rotatable).

On remarquera que ces critères conduisent à des qualités de modélisation différentes. Certains privilégient une bonne précision sur les coefficients du modèle, d'autres assurent une répartition homogène de l'erreur de prévision.

1.6 Plans d'expériences numériques usuels

Nous avons choisi de discuter des plans standards dans ce mémoire car ils ont été développés à l'origine pour des applications de surface de réponse. Parmi ces plans, les plus courants sont les plans factoriels, les plans de Box-Behnken et composites centrés qui sont faciles à générer. Nous avons choisi de discuter aussi des plans de Doehlert qui répondent particulièrement bien à la question du remplissage de l'espace. [8]

1.6.1 Plans factoriels complets

La méthode la plus simple pour répondre à la problématique du remplissage de l'espace consiste à choisir les points sur une grille régulière dans le domaine expérimental. Pour obtenir une grille régulière à k niveaux, il suffit de choisir k valeurs réparties à différents intervalles réguliers pour chaque facteur. Par exemple, dans le carré unité $[0,1] \times [0,1]$ et en choisissant 5 niveaux, on peut considérer la grille de points (voir Figure 1.2) suivante $\{0, 0.25, 0.5, 0.75, 1\} \times \{0, 0.25, 0.5, 0.75, 1\}$.

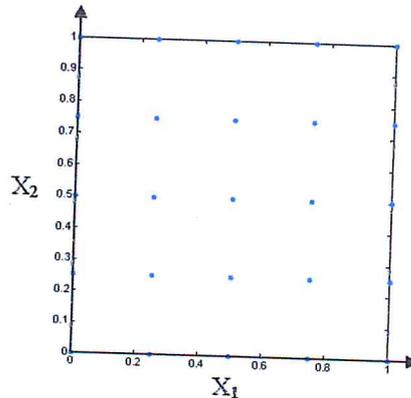


Figure 1.10 Plans composite pour trois facteurs

1.6.2 Plans factoriels fractionnaires:

Le contexte dans lequel nous sommes ne nous autorisant pas à effectuer un nombre importants de simulations, les plans factoriels complets sont inadaptés. Pourtant le principe sur lequel ils reposent reste intéressant. C'est pourquoi, les plans factoriels fractionnaires peuvent constituer une bonne alternative. Le principe consistant à prendre des sous-ensembles de plans factoriels complets permet de réduire considérablement le coût en nombre de simulations (pour plus de détails voir

Myers et Montgomery, 1995). Cependant, les problèmes évoqués en relation avec les plans factoriels au sujet des projections sur les axes factoriels demeurent toujours présents. De nouveaux problèmes d'alignements dus aux alias de la fraction apparaissent.

1.6.3 Plans composites

Une matrice d'expériences composite est une combinaison

- D'une matrice d'expériences factorielle à 2 niveaux, complète (2^d) ou fractionnaire (2^{d-r}), dont les points sont pris parmi les sommets de l'hypercube $[-1,1]^d$ par exemple)
- D'une matrice d'expériences axiale composée de points disposés symétriquement sur chaque axe, à une distance α du centre du domaine
- Et d'un point central pour d facteurs qui permet d'obtenir une information sur la variabilité du phénomène ou tester la validité du modèle (par exemple, dans le cas d'un modèle linéaire de degré 1, tester la présence d'une courbure).

Un composite à faces centrées sur le domaine cubique $[-1,1]^3$ correspond donc à un plan factoriel à 3 niveaux (-1, 0, 1). A noter que ces plans sont adaptés à la démarche « one-at-a-time » qui va imposer des points sur les axes et les plans factoriels 2^d .

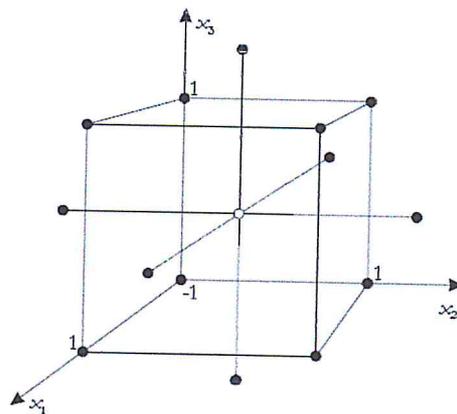


Figure 1.3 -centrale composite k=3.

1.6.4 Plans de Box-Behnken

Un plan de Box-Behnken est un plan dont les variables ne prennent que 3 niveaux ($-b, 0, +b$) si l'on considère pour le domaine expérimental l'hypercube $[-1,1]^d$. Ces plans d'expériences sont composés d'une matrice factorielle à 2 niveaux (2^d points) et de blocs incomplets équilibrés en suivant un arrangement particulier. Ces matrices sont ensuite complétées d'un point au centre du domaine. Ces plans sont une alternative aux plans composites parce qu'ils requièrent seulement 3 niveaux pour chaque facteur (Box et Behnken, 1960) et permettent aussi de modéliser une surface de réponse de degré 2. On peut trouver dans Myers et Montgomery (1995) ou bien aussi dans Dreesbeske, Fine et Saporta (1994) la méthode de construction de ces plans et plus précisément la manière de former les blocs.

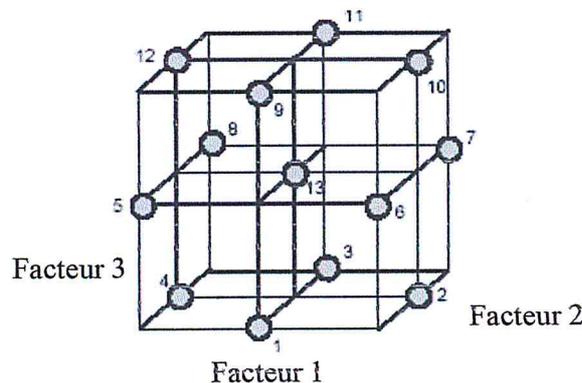


Figure 1.12 Plans Box-Behnken pour trois facteurs

1.6.5 Plans de Doehlert ou réseaux uniformes

Les plans de Doehlert (1970) sont une partie intégrante de la famille des réseaux. La méthode de génération est itérative et consiste à considérer un simplexe initial dans le domaine d'exploration et à lui appliquer des isométries (translations et rotations) à partir d'un de ses sommets (en général, on considère des translations). En itérant ce procédé, on obtient alors une distribution de points particulière (Figure 1.11). Pratiquement, cela revient, pour chaque variable sur $[-1,1]$, à soustraire successivement les coordonnées de chaque point du simplexe initial aux autres.

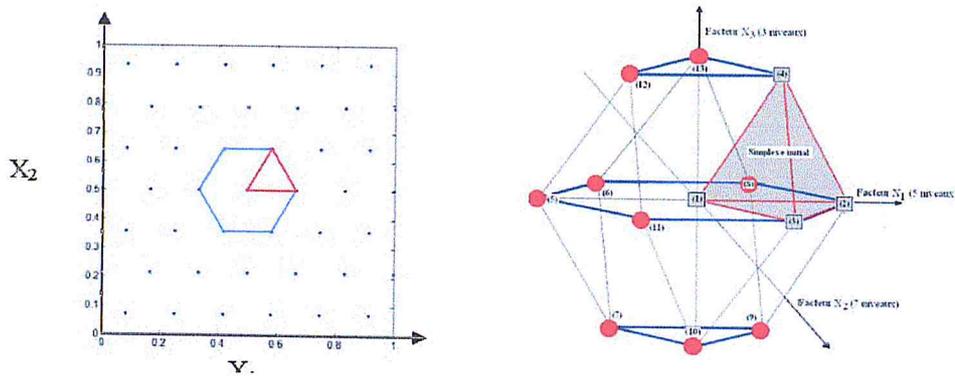


Figure 1.13 Plans Dohler pour deux et trois facteurs

Les plans étudiés auparavant sont généralement de bonne qualité concernant le remplissage de l'espace. Cependant, la plupart de ces méthodes sont soit coûteuses en nombre de simulations ou/et soit une cause de perte d'information dans le contexte des plans d'expériences numériques. C'est essentiellement pour ces raisons que nous avons choisi de n'étudier qu'une infime partie de tous les plans d'expériences classiques connus. Cela ne signifie pas qu'il faut les écarter pour autant des plans d'expériences numériques, le lecteur souhaitant en savoir davantage sur ces plans peut se référer à de nombreux ouvrages, par exemple Myers et Montgomery (1995), Lewis, Mathieu et Phan-Tan-Luu (1999).

1.7 Plans Marginaux

Nous introduisons dans cette section des plans qui, par construction, ont des bonnes propriétés en termes de non redondance, et de non alignement sur certains sous-espaces. Cependant, rien ne garantit qu'ils remplissent bien l'espace. Nous regarderons ici ce qu'il en est.

On définit par ailleurs ce que l'on entend par marges. La notion de marges se réfère à des sous-espaces factoriels. Par exemple, les marges de dimension 1 correspondent aux axes factoriels.[8]

1.7.1 Hypercubes latins

Chaque axe $[0,1]$ du cube unité est découpé en n segments de même longueur

de la façon suivante $\left\{ \left[0, \frac{1}{n} \right], \left[\frac{1}{n}, \frac{2}{n} \right], \dots, \left[\frac{n-1}{n}, 1 \right] \right\}$. En faisant le produit cartésien de ces intervalles, on obtient un maillage de n^d cellules de même taille. Ensuite, on choisit n cellules parmi les n^d possibles de telle manière que les cellules des marges de dimension 1 soient représentées une fois chacune. Enfin, on tire un point au hasard dans chaque cellule présélectionnée.

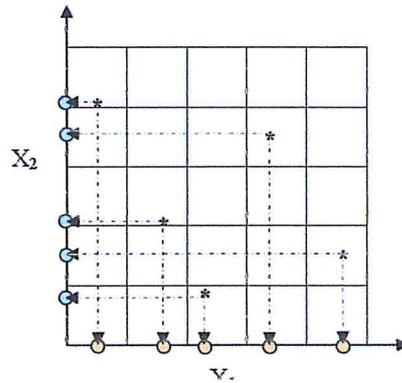


Figure 1.14 Un hyper cube latin à 5 points à 2 dimensions

Remarques

- Nous pouvons choisir de placer le point au centre des cellules pour ainsi supprimer l'aléa de la planification.
- Les points d'un hyper cube latin ont la propriété intéressante d'être uniformément distribués sur les axes factoriels. Cependant, cette propriété n'assure pas l'uniformité des points sur tout le domaine expérimental.

1.7.2 Tableaux orthogonaux

Conceptuellement, les tableaux orthogonaux (Owen, 1992) sont très proches des hyper cubes latins. En effet, ils ont l'avantage d'avoir les bonnes propriétés de projection en grande dimension qu'ont les hyper cubes latins en dimension 1.

Définition 1 : Un tableau orthogonal de force t à q symboles est une matrice de n lignes et d ($> t$) colonnes avec ses éléments prenant q valeurs possibles, telle que toutes les sous matrices de taille $n \times t$ aient q^t lignes apparaissant le même nombre λ fois.

Il est évident que $n = \lambda q^t$. On note un tel tableau orthogonal $OA_\lambda(n,d,q,t)$.

Description : Géométriquement, cela revient à découper les axes du cube unité en q segments égaux afin d'obtenir q^d cellules de même taille. On choisit n cellules de telle façon qu'elles forment un tableau orthogonal défini comme ci-dessus. Cette définition assure que chaque t colonnes de la matrice du plan, soit chaque t -uplet, apparaît exactement λ -fois.

Remarque : Un tableau orthogonal de force 1 est un hyper cube latin. Comme pour les hyper cubes latins, on peut choisir le point aléatoirement dans la cellule mais aussi au centre de la cellule. Dans ce dernier cas, toutes les projections sur les sous espaces de dimensions t conduisent alors à une grille régulière comme on peut le constater sur la Figure 1.13.

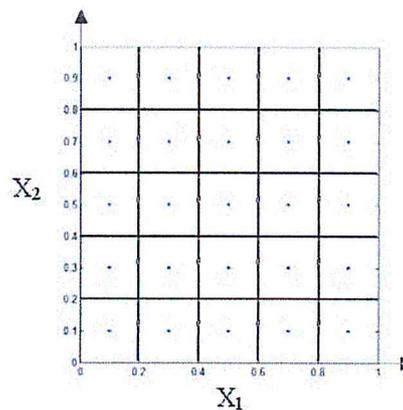


Figure 1.15 montre les points générés par un tableau orthogonal de type OA_1 projetés sur le sous-espace (X_1, X_2)

1.8 Les Suites à faible discrédance

Dans la partie précédente, nous avons discuté de plans dont les points sont bien répartis en projection mais pas forcément dans l'espace. Nous présentons ici des plans dont l'objectif est de bien remplir l'espace. Aussi, nous nous attachons à regarder leurs propriétés en projection. Ces suites de points sont initialement utilisées pour remplacer les suites aléatoires dans la méthode de Monte-Carlo, d'où la dénomination de méthodes de quasi Monte-Carlo. La plupart de ces suites à faible discrédance (Thiérmard, 2000) sont générées par des algorithmes déterministes afin

d'obtenir des points répartis le plus uniformément possible dans le domaine expérimental. On donne ici au lecteur une définition simple et basique de la discrédance pour comprendre en quoi elle répond à nos attentes de remplissage de l'espace.[5]

Définition 2:

La discrédance mesure l'écart entre une distribution de points uniforme et une distribution de points donnée ; en d'autres termes elle permet de mesurer l'irrégularité de la distribution.

Si (dans le cas de la dimension 1) l'on considère la fonction de répartition empirique \hat{F}_n des points x^0, \dots, x^{n-1} , on peut définir la discrédance par :

$$D_n(x) = \sup_{x \in [0,1]} |\hat{F}_n(x) - F_u(x)|$$

avec F_u la fonction de répartition de la loi uniforme sur $[0,1]$.

Remarque:

$D_n(x)$ est la statistique de Kolmogorov-Smirnov permettant de tester l'adéquation à la loi uniforme

Propriété:

Une suite (x_n) est uniformément distribuée si : $\lim_{n \rightarrow \infty} D_n(x) = 0$.

Il existe de nombreux résultats de majoration de la discrédance. Commençons par énoncer le résultat le plus fameux, connu sous le nom d'inégalité de Koksma-Hlawka.

Théorème :(Hlawka 1961)

Si f est une fonction à variation $V(f)$ bornée au sens de Hardy et Krause, alors pour toute suite de points $\{x^0, x^1, \dots, x^{n-1}\} \in [0,1]^d$ on a :

$$\left| \frac{1}{n} \sum_{k=1}^n f(x_k) - \int_{[0,1]^d} f(t) dt \right| \leq V(f) D_n(x)$$

Il est possible de construire des suites pour lesquelles la discrédance est plus faible que celle d'une suite aléatoire. On les appelle des **suites à faible discrédance**. Ces suites ont la particularité de remplir le cube unité uniformément et de manière extrêmement régulière.

Aussi, nous verrons que la complexité du calcul de la discrédance est dépendant de la dimension et devient donc vite inenvisageable en grande dimension. C'est pourquoi, les méthodes de suites à faible discrédance proposées ici semblent intéressantes car elles sont simples à mettre en œuvre et garantissent une faible discrédance.

Une notion importante sur laquelle repose la construction de la plupart de ces suites est la définition de la fonction radicale inverse en base b donnée ci-dessous.

Toutes les suites données ci-après sont définies pour tout n. Nous verrons que la plupart de ces suites sont intéressantes vis-à-vis de l'aspect itératif. Par exemple, lorsque l'on souhaite rajouter q points à un plan de taille n. C'est notamment cet aspect qui va influencer le choix de la suite à utiliser.

Définition 3:

Soit b un entier supérieur à 2. La fonction radicale inverse en base b est donnée par :

$$\phi_b(i) = \frac{p_0}{b} + \frac{p_1}{b^2} + \dots + \frac{p_m}{b^{m+1}}$$

$$i = p_0 + p_1 b + \dots + p_m b^m \quad \text{et} \quad m = \begin{cases} 1 + \lfloor \log_b(i) \rfloor & i \neq 0 \\ 1 & \text{sinon} \end{cases}$$

La suite $C_b = \{x^0, x^1, \dots, x^{n-1}\}$ avec $x^i = \phi_b(i)$ est appelée suite de Van Der Corput en base b.

1.8.1 Suites de Halton

Les suites de Halton sont la version en dimension $d \geq 1$ des suites de Van Der Corput qui quant à elles constituent la version unidimensionnelle de ces suites. L'idée pour générer les suites de Halton consiste à utiliser une base différente pour chaque dimension.

Définition 4:

Une suite de Halton $H_{b_1, \dots, b_d} = \{x^0, x^1, \dots, x^{n-1}\}$ en bases b_1, \dots, b_d est définie par :

$$x^i = (\phi_{b_1}(i), \dots, \phi_{b_d}(i)) \in [0, 1]^d$$

Où b_1, \dots, b_d sont des entiers positifs premiers entre eux deux à deux.

Remarque :

Ces suites ont l'avantage d'être faciles à implémenter et d'avoir un temps d'exécution très faible. Pour passer de $x^i = \phi_b(i)$ à $x^{i+1} = \phi_b(i+1)$ il suffit d'effectuer une addition, en base b , à droite de $1/b$.

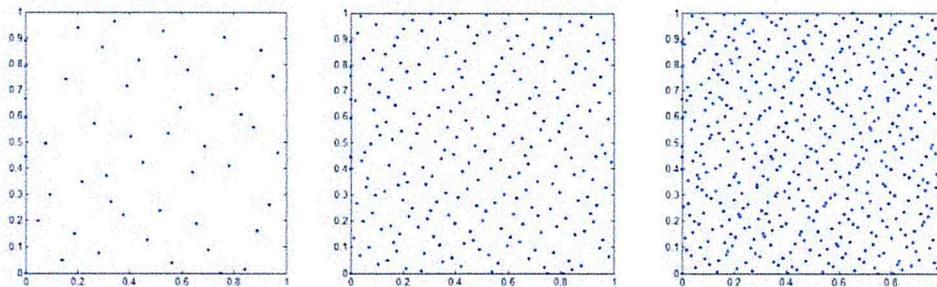


Figure 1.16 De la gauche vers la droite les 50, 250 et 500 premiers points d'une suite de Halton en base 2 et 3

1.8.2 Suites de Hammersley

Une suite de Hammersley en dimension d est construite à partir d'un terme dépendant du nombre de points et d'une suite de Halton en dimension $d-1$.

Définition 5:

Une suite de Hammersley $H_{b_1, \dots, b_{d-1}}^n = \{x^0, x^1, \dots, x^{n-1}\}$ en bases b_1, \dots, b_{d-1} est définie par :

$$x^i = \left(\frac{i}{n}, \phi_{b_1}(i), \dots, \phi_{b_{d-1}}(i) \right) \in I^d$$

Où b_1, \dots, b_{d-1} sont des entiers positifs premiers entre eux deux à deux.

Remarque

Nous prendrons également le soin de prendre les $d-1$ premiers nombre premiers afin de réduire au mieux la discrédance de $H_{b_1, \dots, b_{d-1}}^n$.

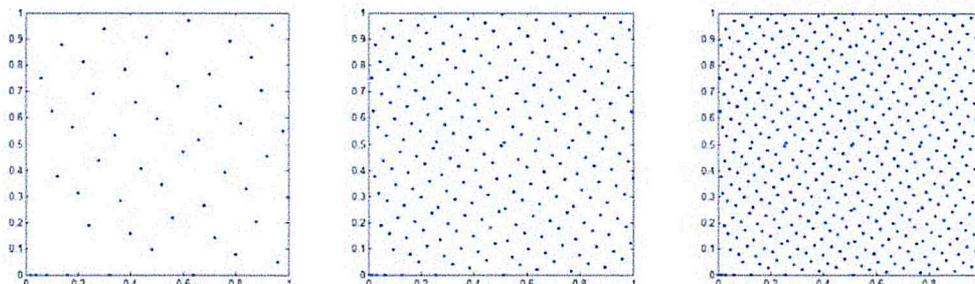


Figure 1.17 De la gauche vers la droite les 50, 250 et 500 premiers points d'une suite de Hammersley en base 2

1.8.3 Suites de Sobol

Elles sont définies à partir de polynômes primitifs sur $Z_2 = [0,1]$. On rappelle à cet effet la définition d'un polynôme primitif.

Définition 6:

Un polynôme $p(t)$ de degré s de la forme $t^s + u_1 t^{s-1} + \dots + u_{s-1} t + u_s$ est primitif sur le corps Z_2 s'il est irréductible sur Z_2 et si le plus petit entier i pour lequel il divise $t^i - 1$ (ou $t^i + 1$) est égal à $2^s - 1$.

Définition 7:

Une suite de Sobol' $S = \{x^0, x^1, \dots, x^{n-1}\}$ en dimension 1 est définie comme suit :

$$x^i = \frac{1}{2^m} \left(\bigoplus_{k=1}^m p_k l_k \right)$$

Où (p_1, \dots, p_m) est la représentation binaire de i , $m = \begin{cases} 1 + \lfloor \log_2(i) \rfloor & i \neq 0 \\ 1 & \text{sinon} \end{cases}$ et où le

Le symbole \oplus désigne simplement l'addition dans Z_2^s .

Les l_k avec $k > s$ sont obtenus en utilisant la relation de récurrence suivante :

$$l_k = 2u_1 l_{k-1} \oplus 2^2 u_2 l_{k-1} \oplus \dots \oplus 2^{s-1} u_{s-1} l_{i-s+1} \oplus (2^s l_{i-s} \oplus l_{i-s})$$

où les u_k sont les coefficients d'un polynôme $t^s + u_1 t^{s-1} + \dots + u_{s-1} t + u_s$ primitif de degré s sur Z_2 et les nombres l_1, \dots, l_s doivent être des entiers impairs tels que $1 \leq l_k \leq 2^k$ pour $k=1, \dots, s$.

Remarque:

A fin de construire une suite de Sobol' en dimension d , il suffit de choisir d polynômes primitifs distincts.

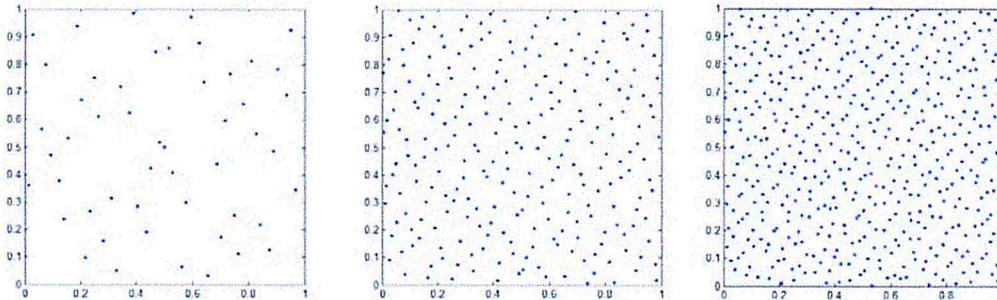


Figure 1.18 les 50, 250 et 500 premiers points d'une suite de Sobol' en dimension 2

1.8.4 Suites de Faure

Les suites de Faure sont définies à partir de la fonction radicale inverse, ϕ_b , et une matrice génératrice de Pascal.

$$C = (c_{kl}) \text{ donnée par } : c_{kl} = \binom{l-1}{k-1} = \begin{cases} \frac{(l-1)!}{(k-1)!(l-k)!} & \text{si } l \geq k \\ 0 & \text{sinon} \end{cases} \quad \forall k, l \in N^*$$

Définition 8:

Soit $b \geq d$, un nombre premier. La suite de Faure $F = \{x^0, x^1, \dots, x^{n-1}\}$ en dimension d est définie comme suit : $x_j^i = \phi_b(C^{j-1}i)$

Où $C^{j-1} = \binom{l-1}{k-1} (j-1)^{l-k} \pmod{b}$, est la matrice génératrice de la j-ième dimension d'une suite de Faure en dimension d.

Remarque: Il est recommandé afin d'obtenir une distribution uniforme de choisir comme base b le plus petit nombre premier supérieur ou égal à d.

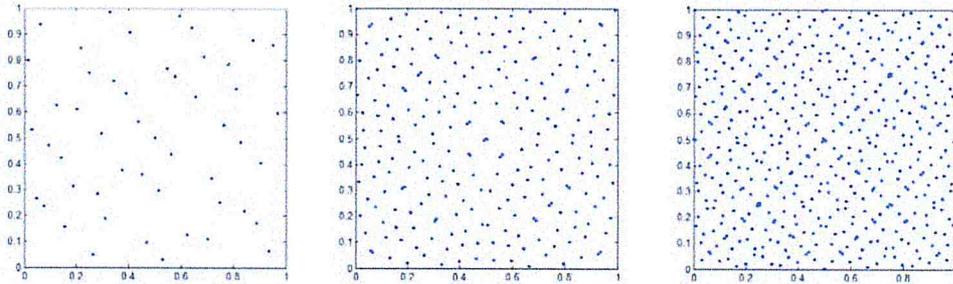


Figure 1.19 les 50, 250 et 500 premiers points d'une suite de Faure en dimension 2

1.8.5 Good Lattice Points

Les good lattice points sont la forme la plus simple des méthodes de réseaux (voir Sloan et Joe, 1994). Les good lattice points peuvent être considérés comme des suites à faible discrédance qui présentent des propriétés analogues à celles des grilles régulières telles que l'espacement régulier entre les points.

P est un good lattice points s'il est défini par:
$$p = \left[\left(\frac{i-0.5}{n} \right), \left(\frac{g_2 i - 0.5}{n} \right), \dots, \left(\frac{g_s i - 0.5}{n} \right) \right]$$
 tq: $i=1, \dots, n$, (g_1, g_2, \dots, g_s) vecteur d'entier.

1.9 Les plans des mélanges

Les expériences de mélanges représentent une classe particulière d'expériences de surface de réponse dans lesquelles le produit analysé est constitué de plusieurs composantes ou éléments. Les plans utilisant ces expériences s'avèrent utiles, car de nombreuses activités de conception et de développement industrielles impliquent des formulations ou des mélanges. Dans de tels cas, la réponse est fonction des proportions des différents ingrédients constitutifs du mélange. Vous pouvez, par exemple, mettre au point une recette de crêpes à base de farine, de levure, de lait, d'oeufs et d'huile, ou développer un insecticide qui allie quatre composantes chimiques.

Dans l'expérience de mélange la plus simple, la réponse (la qualité ou performance du produit mesurée selon certains critères) dépend des proportions relatives des composantes (ingrédients). Les quantités des composantes, mesurées en masse, volume ou autre unité, s'additionnent en un total standard. Dans un plan factoriel, en revanche, la réponse varie en fonction de la quantité de chaque facteur.

Minitab propose trois plans (de mélange centré, en réseaux et sous contraintes) et analyse les données provenant des trois types d'expériences suivants :

- **Mélange** : La réponse est supposée ne dépendre que des proportions des composantes du mélange. Par exemple, la couleur de la peinture ne dépend que des pigments utilisés.
- **Variable de procédé du mélange** : La réponse est supposée dépendre des proportions relatives des composantes et des variables de procédé (dans une expérience, celles-ci sont les facteurs qui ne font pas partie du mélange, mais qui peuvent influencer sur ses propriétés). Par exemple, le goût d'un gâteau dépend des proportions d'ingrédients dans la préparation pour gâteaux ainsi que de la durée et de la température de cuisson.
- **Quantité de mélange** : La réponse est supposée dépendre des proportions des composantes et de la quantité du mélange. Par exemple, le rendement d'une récolte dépend des proportions des ingrédients d'un insecticide et de la quantité d'insecticide appliquée.

1.10. Récapitulatif

Le tableau ci-dessous permet de récapituler les propriétés des plans étudiés au sein de ce chapitre. Chaque type de plan est critiqué en regardant si :

- Les distributions obtenues permettent ou non de remplir l'hyper cube uniformément,
- Les distributions, notamment en projection dans le carré unité, suivent des alignements,
- Les distributions sont séquentielles c'est-à-dire s'il est possible de rajouter des points à cette distribution sans pour autant détruire l'uniformité initiale,
- A partir des distributions obtenues, les distributions en dimension inférieure ont une meilleure uniformité.

Tableau 1.2 Récapitulatif des propriétés de certains plans d'expériences

	Remplissage de l'espace	Motif	Séquentialité	Réduction de Dimension
Aléatoires	Non : lacunes.	Non	Oui	Oui
Doehlert	Oui	Oui	Oui	Non
Grilles régulières	Oui	Oui	Non	Non
Hypercubes latins	Oui, en version centrée. Sinon, lacunes.		Non	Oui, en dimension 1
Hypercubes latins de Tang	Oui	Non	Non	Oui
Tableaux orthogonaux	Oui, en version centrée. Sinon, lacunes.	Oui	Non	Oui, en dimension égale à la force du tableau.
Good lattice points	Oui, si les paramètres sont bien choisis	Oui	Oui	Non
Halton	Oui	Oui, en Dimension élevée.	Oui	Non
Hammersley	Oui	Oui, en Dimension élevée.	Oui	Non
Faure	Oui	Oui	Oui	Non
Niederreiter	Oui	Non	Oui	Non
Sobol'	Oui	Non, si les états initiaux sont bien choisis	Oui	Non

1.11 Présentation des méthodes d'évaluation de la qualité d'une structure

L'évaluation de la qualité de la structure d'un ensemble de points issu d'une base de données ou d'un plan d'expériences nécessite l'utilisation de critères quantitatifs. Il existe de nombreux critères qui permettent, entre autres, d'évaluer l'uniformité d'une répartition de points tels que la discrédance, le critère Audze-Eglais, les critères basés sur les distances, ... En effet, la propriété d'uniformité semble garante d'une bonne représentativité. Dans le cas de l'étude de structures en grande dimension, nous n'avons retenu que les critères basés sur les distances euclidiennes entre points, qui même en grande dimension sont faciles et rapides à

calculer. On note $dist(x^i, x^k)$ la distance euclidienne entre deux points x^i et x^k telle que :

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

1.11.1 Critère Mindist:

Johnson et al. (1990) ont introduit les distances maximin et minimax afin de construire des plans répondant à la question de remplissage de l'espace.

Le critère de distance Maximin, noté Mindist, définit la plus petite distance entre deux points quelconques d'une distribution :

$$Mindis = \min_{\substack{x^i \in x \\ x^k \in x \\ i \neq k}} \min dist(x^i, x^k)$$

Avec $x = \{x^0, x^1, \dots, x^{n-1}\} \in [0,1]^d$ une suite de n points en d dimensions.

Maximiser le critère MinDist revient à considérer le critère de distance **maximin**. Une valeur élevée du critère Mindist sera synonyme d'une bonne répartition des points dans l'espace des variables. A contrario, si une distribution de points présente une valeur Mindist faible, cela signifie qu'il y aura au moins deux points très proches ce qui révèle une zone plus dense dans le domaine que nous pourrions qualifier d'amas. En présence d'amas, la répartition des points dans le domaine ne sera plus idéale au sens de l'uniformité.

1.11.2 Critère Moydist:

Le critère MoyMin définit la moyenne des distances minimales, et renseigne sur l'ensemble des distances minimales entre les points :

$$MoyMin = \frac{1}{n} \sum_{i=1}^n \min_{i \neq k} dist(x^i, x^k)$$

Un ensemble de N points pourra être qualifié d'uniforme si la valeur MoyMin est grande et proche de la valeur Mindist.

Exemple : Pour une distribution de N points dans un espace en D dimensions, l'analyse des valeurs Mindist et MoyMin renseigne sur la qualité de cette distribution au sens de l'uniformité. Par exemple, une valeur Mindist faible et une valeur MoyMin

élevée, signifient qu'au moins deux points sont très proches mais que les autres points sont bien répartis dans l'espace des variables. Si les valeurs *Mindist* et *MoyMin* sont équivalentes et faibles, cela signifie que les plus petites distances entre points sont équivalentes c'est-à-dire que tous les points sont proches et forment un seul amas. A contrario, si une distribution présente des valeurs *Mindist* et *MoyMin* élevées, cette distribution sera plus uniforme (figure 1.16).

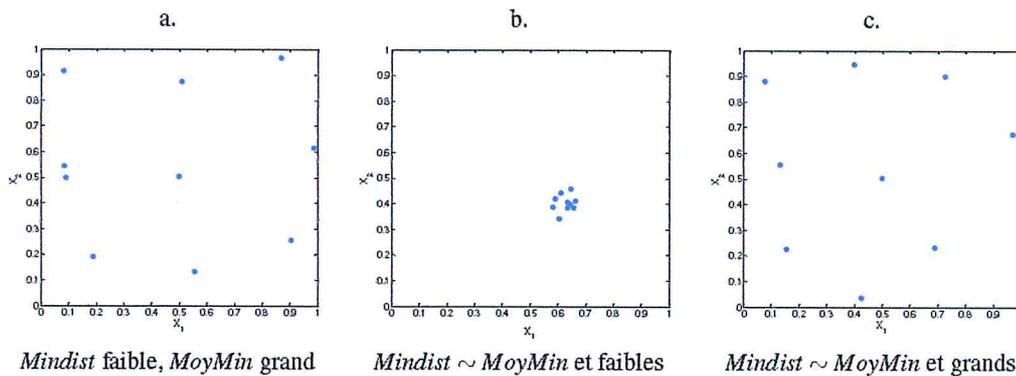


Figure 1.20 Comparaison de trois distributions de points. a) distribution uniforme avec deux points très proches, b) tous les points sont proches et constituent un seul amas, c) distribution uniforme

1.11.3 Critère de recouvrement

Le critère Coverage proposé par Gunzburger, mesure le recouvrement de l'espace par les points et permet de quantifier l'homogénéité de la répartition des points:

Soit $\{x^0, x^1, \dots, x^{n-1}\} \in [0,1]^d$ une suite de n points en d dimensions. On définit la mesure de recouvrement λ par :

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

Avec $\gamma_i = \min_{i \neq j} \text{dist}(x^i, x^j)$ désignant la distance minimale entre le point x^i et les

autres points de la suite et $\bar{\gamma} = \frac{1}{n} \sum_{i=1}^n \gamma_i$ la moyenne des γ_i . Si les points sont sur une

grille régulière, on a: $\gamma_1 = \gamma_2 = \dots = \gamma_n$ et donc $\lambda = 0$. Ainsi, plus λ est faible, et plus les points sont proches de ceux d'une grille régulière. Cette expression fait apparaître explicitement le coefficient de variation de l'échantillon des γ_i , à savoir le rapport de l'écart-type avec la moyenne.

1.11.4 Critère de rapport de distances

Soit $\{x^0, x^1, \dots, x^{n-1}\} \in [0,1]^d$ une suite 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 $\gamma_i = \min_{i \neq j} \text{dist}(x^i, x^j)$

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.

1.11.5 Critère discrédance

Dans le cas général, le calcul de la discrédance est une tache réputée très difficile. Il a pourtant été montré (voir Niederreiter [Nie72]) que le problème est discrétisable et résoluble en un nombre fini d'étapes. Malheureusement, la complexité des différents algorithmes connus croit de manière exponentielle avec la dimension. Par ailleurs, la difficulté du problème est telle que l'on se contente souvent de bornes inférieures et supérieures sur la discrédance. En revanche, il se trouve que la discrédance carrée moyenne $(D_n^*(x))^2$ se laisse facilement calculer dans n'importe quelle dimension :

Théorème 1.1 (Warnock) Pour une séquence $\{x^0, x^1, \dots, x^{n-1}\} \in [0,1]^d$, on a :

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

et

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

La formule de Warnock montre que la discrédance carrée moyenne d'une séquence de n points peut être calculée en $O(n^2)$ opérations. Retravaillant le dernier terme de cette expression, Heinrich] a obtenu un algorithme en $O(n(\log n)^d)$. Hélas,

comme le laissait déjà supposer sa complexité théorique, les expériences numériques de Heinrich ont révélé qu'en dimension élevée ($d > 8$), sa méthode n'est pas plus performante que l'approche directe. D'autre part, se fondant sur une analyse de complexité détaillée, Matousek a montré que l'algorithme de Heinrich n'est réellement plus efficace qu'une application directe de la formule de Warnock que pour des séquences comprenant au moins 2^{2d} points.

CHAPITRE 2

SIMULATION D'UN PROCESSUS PONCTUEL MARQUE PAR LA METHODE DE MONTE CARLO PAR CHAINE DE MARKOV (M.C.M.C)

Dans ce chapitre, nous présentons les principales définitions et résultats relatifs à la théorie des processus spatiaux utilisés par la suite dans ce mémoire. Ainsi nous présentons les processus ponctuels et leur extension en processus objet. Puis les chaînes de Markov utilisées pour la simulation de ces processus. Enfin nous étudierons les algorithmes de simulation pour les chaînes de Markov.

2.1 Processus ponctuels

Les processus ponctuels (distributions spatiales aléatoires des points) sont le fondement de la géométrie stochastique. Des exemples de l'utilisation de ces processus stochastiques peuvent être observés dans des domaines aussi variés que la biologie, la mécanique, la physique des particules, l'ingénierie au sens large, etc...

Soit (χ, d) un espace métrique, χ définissant l'espace muni d'une métrique d (en général on utilise $\chi = \mathbb{R}^n$, d étant définie comme une distance euclidienne). Une configuration est un ensemble dénombrable non ordonné de points de χ :

$$x = \{x_1, \dots, x_j, \dots, x_n\}, n \in \mathbb{N}$$

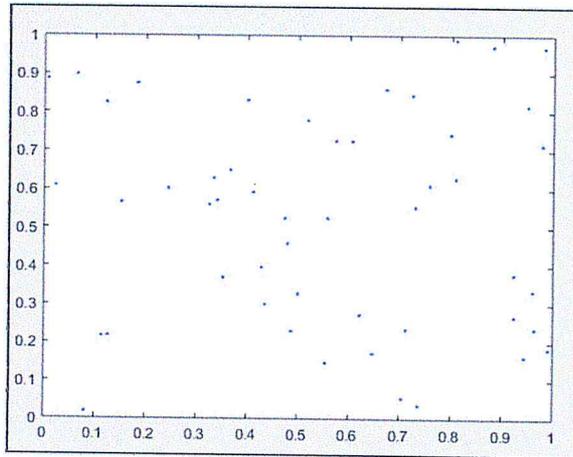


Figure 2.1 : une configuration de points dans $[0, 1]^2$

Par la suite, on s'intéressera aux configurations x dites localement finie c'est à dire qui place dans tout borélien borné $A \subseteq \chi$ un nombre $\aleph_x(A)$ fini de points et simple (tous les points $x_i \in x$ sont distincts). Définissons un espace probabilisé (Ω, A, \mathcal{P}) constitué d'un ensemble Ω , d'une σ -algèbre A qui contient les sous ensembles de Ω et d'une mesure de probabilité \mathcal{P} définie sur cette σ -algèbre telle que $\mathcal{P}(\Omega) = 1$.

Un processus ponctuel est une application X mesurable définie sur cet espace probabilisé (Ω, A, \mathcal{P}) vers \aleph_x la famille des configurations de points de χ localement finies telle que pour tout borélien borné $A \subseteq \chi$, le nombre des points en A soit une variable aléatoire (finie). [2]

Nous présentons des modèles plus simples et plus utilisés de processus ponctuels :

- 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 :

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

Avec $s(x) = \sum_{i < j} 1_{|x_i - x_j| \leq R}$ et $0 < \gamma < 1$ un coefficient de répulsion, k la constante de normalisation, et $s(x)$ est le nombre d'interaction entre chaque paire de point. [10]

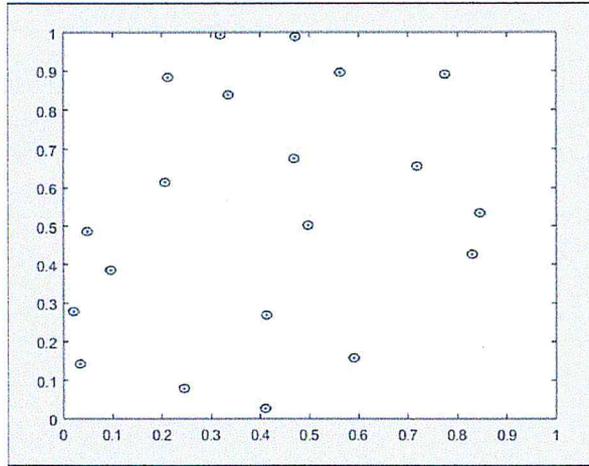


Figure 2.2 : Plan de Strauss à 20 points en dimension 2 et $\gamma = 0.2$ et de $rayon = 0.01$.

- Le modèle de Gibbs s'écrit sous la forme suivante :

$$\pi(x) = k \exp^{-U(x)}$$

Avec k , la constante de normalisation et $U(x)$ est la fonction de potentiel. Pour $\Omega = [0,1]^d$ nous parlons du modèle de Besag [1974] et Guyon [1995], où Φ le potentiel d'interaction et la fonction du potentiel $U(x)$ est définie comme suit :

$$U(x) = \sum_{i \in \Omega} \Phi_i(x_i) + \sum_{\langle i,j \rangle} \Phi_{ij}(x_i, x_j)$$

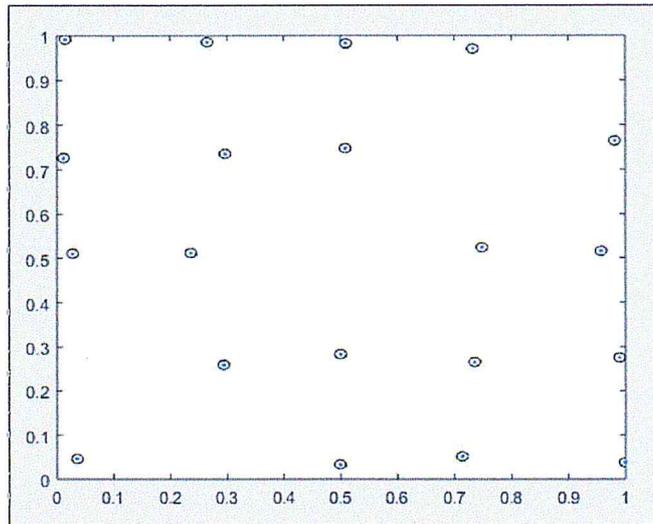


Figure 2.3 : plan de Gibbs à 20 points en dimension 2 et de $rayon = 0.01$.

2.2 Processus ponctuels marqués

Les processus ponctuels sont les exemples les plus simples de la géométrie stochastique, mais on a souvent besoin de simuler des processus impliquant des objets plus complexes.

Pour décrire des objets au lieu de points, il est simplement nécessaire d'ajouter à chaque point une « marque » qui décrit les propriétés de l'objet (sa dimension, son orientation, sa forme, etc.). Ce processus est alors nommé « processus de points marqués » ou « processus objet » (Daley et Vere-Jones 2002).

Un processus objet est une séquence aléatoire $X = \{P_n, M_n\}$ formé d'un processus de points X_n défini sur un sous-espace χ de \mathbb{R}^d et des marques M_n correspondantes à chaque X_n dans l'espace des marques M . Si nous cherchons à simplifier la définition, nous pouvons considérer le processus objet X avec les points dans χ et les marques dans M comme un simple processus ponctuel sur $\chi \times M$. Pourtant, en pratique, il est souvent plus intéressant de séparer les points et les marques pour pouvoir définir plus simplement des translations dans l'espace des points tout en laissant inchangé les marques attachées à chaque objet où, au contraire pour avoir la possibilité de changer la forme de l'objet sans nécessairement le déplacer. [9]

2.2.1 Quelques modèles

Les densités des processus objets couramment utilisés sont de la forme :

$$\pi(x) = a(n(x))g(x)$$

où $n(x)$ le nombre des objets dans la configuration x , avec $a(x)$ et $g(x)$, deux fonctions mesurables positives.

Dans le cas d'un processus de poisson $a(n) = \beta^n$ et $g(x) = 1$
 $g(x)$, modélise les interactions entre objets et $a(x)$ est l'intensité du processus i.e. le nombre moyen d'objets de x . [4]

Le processus de Strauss possède en fait une particularité par rapport à la définition générale des processus avec interaction par paires d'objets, dans le processus de Strauss, deux objets u et v n'interagissent que s'ils vérifient une relation de voisinage $u \sim v$ symétrique sur U , $g(x) = \gamma^{1_{u \sim v}}$. Cette relation est à priori quelconque, mais on la définit généralement par $u \sim v \leftrightarrow d(u, v) \leq R$, d est la distance entre deux objets. Sa densité π s'exprime alors par :

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

Où, $\beta > 0$ et $0 < \gamma < 1$ et $s(x) = \sum_{i < j} 1_{\|x_i - x_j\| \leq R}$ désigne le nombre de paires d'objets qui sont à une distance inférieure ou égale à R l'un de l'autre.

- pour $\gamma = 1$, les interactions confèrent au modèle le comportement d'un processus de Poisson avec une intensité $\beta\lambda(\cdot)$
- pour $0 < \gamma < 1$, une force de répulsion est définie entre les objets
- pour $\gamma = 0$ le terme d'interaction est dit « hard-core » et interdit toute interaction entre les objets.
- pour $\gamma > 1$, le processus devient attractif.

2.2.2 Processus de Markov

Ripley et Kelly ont introduit les processus ponctuels de Markov qui sont une généralisation des processus de Strauss. Leur définition équivaut en effet à dire qu'un processus X est de Markov si et seulement si sa densité est de la forme :

$$\pi(x) = \prod_{y \subseteq x} \Phi(y)$$

Où Φ est une fonction positive telle que $\Phi(y) = 1$ s'il existe u et v dans y tels que $u \not\sim v$ (n'existe pas une relation entre u et v). [10]

Deux exemples des processus Markovien objets aux plus proches voisins sont introduits par Baddeley et Moller qui sont importants pour la modélisation des phénomènes répulsifs. Ces deux processus sont présentés dans le cas particulier où les objets sont des disques de rayon R fixé.

- Le premier processus à interaction de connexité [Haggstrom, Lieshout et Moller 1999, Kendall 1999], de densité :

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

Où $c(x)$ est le nombre de cliques maximales dans x pour la relation \sim_x (où de façon équivalente le nombre de composante connexe dans x pour la relation \sim).

- Le second processus ponctuel à interaction d'aire [Baddeley et Lieshout 1995], de densité :

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

Où $a(x)$ est l'aire de $B(x) = \bigcup_{x_i \in x} B(x_i, \frac{r}{2}) \cap \mathcal{X}$.

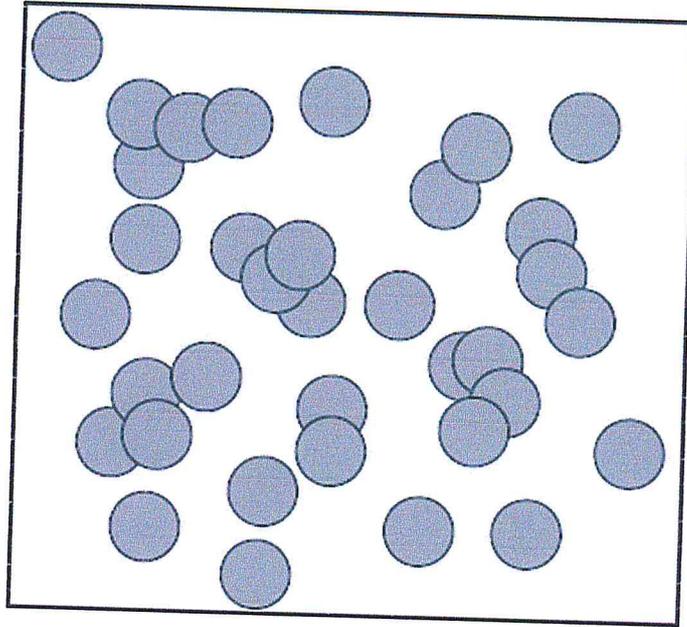


Figure 2.4 : Une réalisation de l'ensemble $B(x)$ permettant de définir l'aire $a(x)$ et le nombre de composantes connexes $c(x)$

2.3 Simulation des processus ponctuels et convergence

Pour que le processus objet puisse atteindre la solution optimale, une solution classique consiste à définir une chaîne de Markov de type Monte Carlo (MCMC). Le but de ces méthodes est de construire une chaîne de configurations $\{x_0, x_1, \dots, x_n\}$ qui converge vers la distribution recherchée π . De nombreuses méthodes ont été proposées pour la simulation des chaînes de Markov (l'algorithme de Metropolis Hastings, l'échantillonneur de Gibbs et l'algorithme de naissance-mort). [2]

Une chaîne de Markov $(x_n)_{n \geq 0}$ sur Ω , de loi initiale π_0 , de probabilités de transition $(P(x, y))_{x, y}$ est une suite de variables aléatoires à valeurs dans Ω telle que :

- pour tout $x \in \Omega$, $P(X_0 = x) = \pi_0$
- pour tout $n \geq 1$ et pour tout $(n + 1)$ -uplet $(x_0, x_1, \dots, x_n) \in \Omega^{n+1}$, on a :

$$P(x_0 = x_{[0]}, x_1 = x_{[1]}, \dots, x_n = x_{[n]}) = \pi_0(x_{[0]})P(x_{[0]}, x_{[1]})P(x_{[1]}, x_{[2]}) \dots P(x_{[n-1]}, x_{[n]}).$$

En particulier, on aura $P(x, y) = P(x_1 = y | x_0 = x) = P(x_{n+1} = y | x_n = x)$. On notera par la suite $P^{(n)}(x, y) = P(x_n = y | x_0 = x) = P(x_{n+k} = y | x_k = x)$, c'est-à-dire la probabilité de passer de l'état x à l'état y en n étapes. [6]

Une chaîne $\{x_0, x_1, \dots, x_n\}$ de configurations aléatoires forme une chaîne de Markov lorsque la prédiction du futur (la configuration suivante) à partir de la configuration présente n'est pas rendue plus précise par des éléments d'information supplémentaires concernant le passé de la chaîne :

$$P(x_{n+1} \in A | x_0, x_1, \dots, x_n) = P(x_{n+1} \in A | x_n) \quad \forall A \in \mathcal{B}$$

D'un point de vue informatique, une telle chaîne présente bien entendu l'avantage de rendre inutile le rappel de l'ensemble des configurations antérieures puisqu'elle utilise uniquement l'état actuel pour générer une nouvelle configuration. La génération de la nouvelle configuration réclame la définition d'un noyau de transition : un noyau de transition est une fonction $P : \Omega \times \mathcal{B} \rightarrow [0,1]$ tel que :

- $\forall A \in \mathcal{B}$, la fonction $P(\cdot, A)$ est mesurable
- $\forall X \in \Omega$, la fonction $P(X, \cdot)$ est une mesure de probabilité

La propriété la plus importante d'une chaîne Markov est la stationnarité. À une chaîne stationnaire correspond une distribution stationnaire π vers laquelle elle converge.

Nous présentons quelques résultats théoriques qui caractérisent la convergence des méthodes de Monte Carlo par Chaîne de Markov:

Invariance [10]

Une loi π est invariante pour la chaîne de Markov si :

$$\pi(A) = \int P(x, A) \pi(dx)$$

Cette condition est nécessaire pour obtenir la convergence de la chaîne vers π .

Réversibilité [10]

La chaîne est réversible pour π si le noyau de transition P vérifie :

$$\int_B P(x, A) \pi(dx) = \int_A P(x, B) \pi(dx) \quad \forall A, B \in \mathcal{B}$$

Cette condition implique l'invariance pour π , et signifie que sous la distribution stationnaire π la probabilité de passer de A à B est la même que de passer de B à A . La plupart des algorithmes de simulation sont en réalité construits pour produire des chaînes de Markov réversibles.

Irréductible [10]

La chaîne est dite π -irréductible si :

$$\forall x \in \Omega \text{ et } \forall A \subset \mathcal{B} \quad \pi(A) > 0 \rightarrow \exists t/P^t(x, A) > 0$$

Cela signifie que la chaîne a une probabilité non nulle d'atteindre en temps fini tout ensemble π -probable. Cette condition est clairement nécessaire pour que la chaîne converge en distribution vers π avec n'importe quelle condition initiale.

Apériodique [10]

L'apériodicité assure que les déplacements entre états n'ont pas trop de contraintes. Formellement, la chaîne est apériodique s'il n'existe pas d'ensembles

$$\{A_0, \dots, A_{d-1}\} \in \mathcal{B} \text{ avec } d \geq 2 \text{ tq } P(x, A_{i+1[d]}) = 1 \quad \forall x \in A_i$$

Primitive [11]

P est un noyau primitif (ou régulier) si il existe $k \geq 1$ tel que P^k ait tous ses termes strictement positifs.

Ergodicité uniforme [12]

La chaîne $(X_k)_{k \geq 0}$ de Markov est uniformément ergodique si $\|vK^r - \pi\| \leq M$

avec : M une constante.

v la loi initiale et π la loi de distribution finale.

K^r la transition de la chaîne.

Théorème 2.1 [3]

Pour un noyau P irréductible sur Ω fini, il y a équivalence entre

(1) P est apériodique

(2) Il existe un entier r tel que $P^n(x, y) > 0$ pour tous x et y et tous $n \geq r$.

Théorème 2.2 [1]

Supposons que P est π -irréductible et π -invariante, alors P est récurrente positive et π est l'unique loi invariante de P .

Définition [7]

Le coefficient de contraction [Dobrushin 1956] défini comme suit est utile vers la fin.

$$C(K) = \frac{1}{2} \max_{x,y} \|P(x, \cdot) - P(y, \cdot)\| \text{ avec } 0 < C(P) < 1$$

Lemme 2.1 [7]

Soient ν et π deux distributions, P et Q deux noyaux de transition. Alors

$$\|\nu P - \pi\| \leq \|\nu - \pi\| C(P) \text{ et } C(PQ) \leq C(P)C(Q)$$

En particulier,

$$\|\pi P - \nu P\| \leq \|\pi - \nu\| \text{ et } \|\pi P - \nu P\| \leq 2C(P)$$

2.4 Les algorithmes de simulation

Pour simuler une configuration d'objets par un processus ponctuel marqué, nous allons construire une chaîne de type Monte Carlo. Il existe plusieurs algorithmes pour la simulation d'une telle chaîne, nous allons en présenter quelques uns dans cette section. Nous pouvons distinguer deux types d'algorithmes pour échantillonner les chaînes MCMC : selon le type d'application, il est en effet possible d'utiliser des chaînes en temps continu où des processus discrets utilisés préférentiellement pour simuler des processus spatiaux (Winkler 2003).[11]

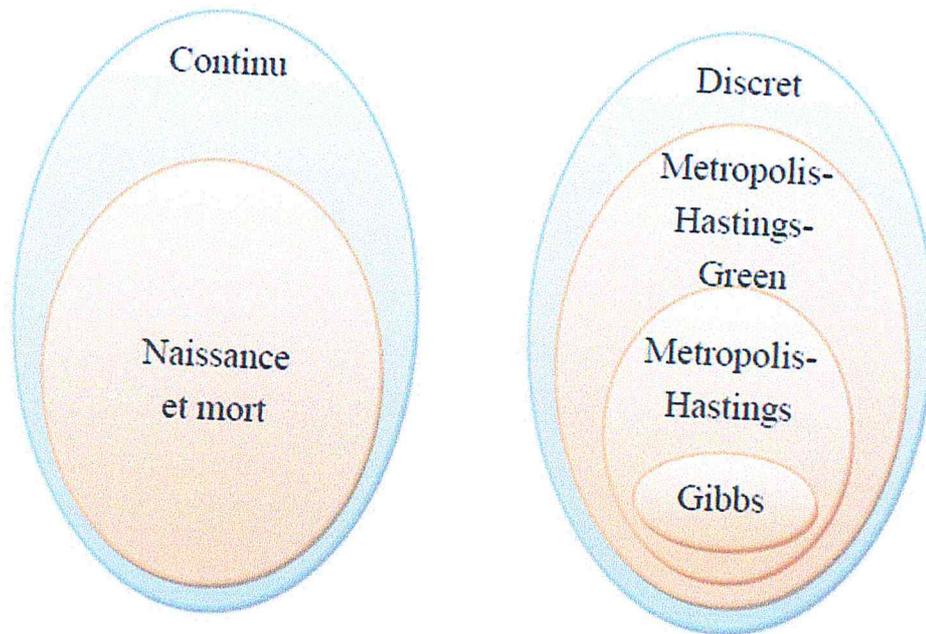


Figure 2.5 : les familles d'algorithmes pour la simulation des chaînes MCMC

- **Les chaînes en temps continu**

La méthode la plus importante est la méthode de naissance-mort (Preston 1976, Ripley et Kelly 1977).

- **Les chaînes discrètes**

Ces méthodes se basent en général sur la méthode Metropolis-Hastings (Hastings 1970). Cet algorithme a été généralisé à travers l'algorithme Metropolis-Hastings-Green (Green 1995). Un autre algorithme classique permettant de simuler les processus discrets également basé sur l'algorithme Metropolis-Hastings est l'échantillonneur de Gibbs, décrit dans (Geman et Geman 1984).

2.4.1 L'algorithme de type Metropolis-Hastings (MH)

L'algorithme Metropolis a été introduit par (Metropolis, et al. 1953) et généralisé par Hastings (1970), il a été adapté pour le cas des processus spatiaux par (Geyer et Møller 1994). L'idée de base de l'algorithme Metropolis-Hastings est de proposer un nouvel état par une petite perturbation de l'état actuel, puis de l'accepter ou le refuser. Le taux d'acceptation $R(x, y)$, représente la probabilité de transition de la chaîne d'un état actuel x dans l'état y . L'algorithme proposé par Geyer et Møller construit une chaîne de Markov qui peut explorer toutes les configurations de l'espace χ . Les perturbations proposées sont la naissance et la mort (ajout ou suppression d'un élément de la configuration actuelle). Il s'écrit comme suite :

A partir d'un état $x_t = x$:

- **Naissance** : Avec une probabilité $\frac{1}{2}$ choisir un nouvel objet $\xi \in \chi$ et proposer de l'ajouter à la configuration actuelle : $y = x \cup \xi$. Le taux d'acceptation sera :

$$R(x, y) = \frac{\pi(y)v(x)}{\pi(x)n(y)}$$

- **Mort** : avec une probabilité $\frac{1}{2}$ choisir un objet de la configuration actuelle $\xi \in x$ et proposer de l'enlever : $y = x \setminus \xi$ et calculer le taux d'acceptation :

$$R(x, y) = \frac{\pi(y)n(x)}{\pi(x)v(\chi)}$$

- La configuration y sera acceptée avec la probabilité $a = \min(1, R)$

Algorithme 2.1 : Metropolis Hastings.

2.4.2 L'Algorithme de Metropolis-Hastings-Green (MHG)

Cette extension de l'algorithme Metropolis-Hastings proposée par (Green 1995) construit aussi une chaîne Monte Carlo à sauts réversibles, mais propose des perturbations plus complexes que la naissance et mort présentées dans la section précédente, Ceci se fait en 2 étapes :

- Noyau de transition de changement : on commence par proposer un changement $x \rightarrow y$ selon une transition $Q(x, \cdot)$.
- Probabilité d'acceptation du changement : ensuite, on accepte ce changement avec la probabilité $a(x, y)$ où $a: \Omega \times \Omega \rightarrow]0, 1]$.

En utilisant un mélange de noyaux simples de transition $Q(x, \cdot) = \sum_m p_m Q_m(x, \cdot)$. Une transition dans la chaîne utilise le noyau combiné Q , les noyaux Q_m sont les noyaux de base. La probabilité de choisir un noyau simple Q_m est p_m . Pour construire une chaîne de Markov, l'algorithme MHG introduit une mesure symétrique $q(x, \cdot)$, Il s'écrit comme suite :

A partir d'un état $x_t = x$:

- Choisir un noyau de proposition Q_m avec une probabilité p_m .
- Générer $y \sim Q(x, \cdot)$
- Calculer le rapport d'acceptation

$$R(x, y) = \frac{\pi(y)}{\pi(x)}$$

- acceptée l'état y avec la probabilité $a = \min(1, R)$.
- recommencer avec l'état actuel.

Algorithme 2.2 : Metropolis Hastings Green.

2.4.3 L'échantillonneur Gibbs

L'échantillonneur Gibbs est un cas spécial de l'algorithme Metropolis-Hastings. La différence entre les deux algorithmes est que l'échantillonneur Gibbs accepte tous les nouveaux états, mais ces états sont générés selon la densité vers laquelle on souhaite converger, contrairement à l'algorithme Metropolis qui génère les nouveaux états selon une

densité auxiliaire uniforme sur l'ensemble des états. L'algorithme Gibbs à partir d'une configuration $x = \{x_1, \dots, x_j, \dots, x_n\}$ est donné par :

- Choisir un objet à modifier $x_j \in x$
- Générer une nouvelle valeur x'_j selon $\pi(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$
- Recommencer avec la nouvelle configuration

$$x = \{x_1, \dots, x'_j, \dots, x_n\}$$

Algorithme 2.3 : Echantillonneur de Gibbs



CHAPITRE 3

GENERATION DE NOUVEAU PLAN A PARTIR DES PROCESSUS PONCTUELS MARQUES

On présente dans ce chapitre, une méthode de construction de plans d'expériences numériques basée sur la méthode de Monte Carlo par chaîne de Markov (MCMC). Ces plans respectent deux critères (sur la distribution des points et sur une certaine caractérisation des marques de ces points). Ainsi, l'utilisation de l'algorithme de Metropolis Hastings Green pour la simulation de la loi π du processus marqué et la justification de sa convergence.

3.1 Plan marqué de Gibbs

Les processus marqués de Gibbs sont fondés sur la densité de Gibbs utilisée en physique statistique. Ces processus ponctuels caractérisés par une certaine marque, sont définis par une densité qui peut être écrite sous une forme énergétique et incorporant des potentiels d'interaction. La densité est donnée par:

$$\pi(x) = k\beta^{n(x)}\exp^{-U(x)} \quad (3.1)$$

Avec

- k la constante de normalisation
- $n(x)$ le nombre d'objets dans la configuration x (voir paragraphe 3.2)
- $\beta \geq 0$
- $U(x)$ désigne le potentiel d'énergie :

$$U(x) = \eta \sum_{i < j} \varphi(\|x_i - x_j\|) + \sum_{k=1}^n \eta_k \sum_{i < j} \varphi_k(|x_i^k - x_j^k|)$$

Où $\eta = -\log(\gamma)$, $\eta_k = -\log(\gamma_k)$, $k = 1, \dots, n$ 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 suivantes :

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

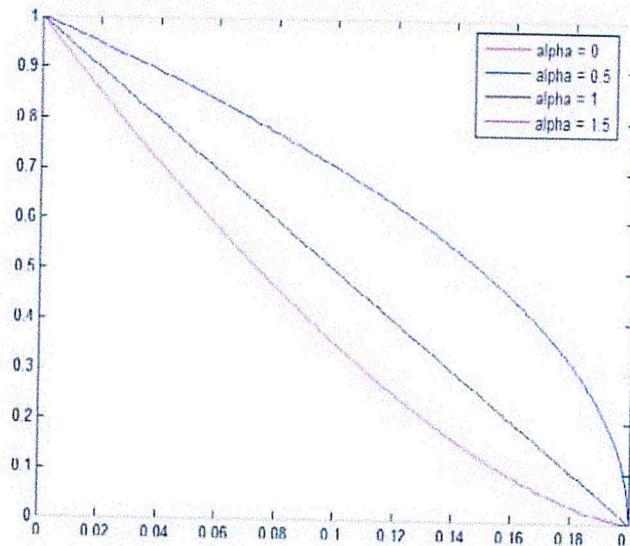


Figure 3.1 La fonction de potentiel puissance pour différents α et $R=0.2$.

3.2 Le choix des marques

3.2.1 Fonction de variance de prédiction

L'erreur commise sur les réponses mesurées dépend de la nature de l'expérimentation, du degré de précision de la technologie employée, du soin, de l'habileté de l'expérimentateur, et de bien d'autres facteurs dont l'expérimentateur est responsable. Ces facteurs ne dépendent pas de la théorie des plans d'expériences mais de la pratique expérimentale [5]. Pour séparer cette partie expérimentale de celle qui dépend de la théorie, on introduit la fonction de variance de prédiction $d^2(\hat{y}_u)$:

$$d^2(\hat{y}_u) = x_u^T (X^T X)^{-1} x_u$$

En prenant la racine carrée de la fonction de variance, on obtient la fonction d'erreur de prédiction :

$$d(\hat{y}_u) = \left[x_u^T (X^T X)^{-1} x_u \right]^{\frac{1}{2}}$$

On peut tracer les courbes d'égale erreur de prédiction dans le domaine d'étude. Il est intéressant de constater que la fonction d'erreur de prédiction ne dépend pas des résultats des expériences c'est à dire des valeurs des réponses mesurées. Elle dépend essentiellement de l'emplacement des points expérimentaux dans le domaine d'étude et du modèle postulé. On

peut donc savoir, avant de commencer l'expérimentation, comment la précision des réponses prédites sera affectée par le choix de l'emplacement des points expérimentaux et par celui du modèle. En faisant apparaître les écarts-types et la fonction de la variance de prédiction, la relation peut s'écrire :

$$\text{var}(\hat{y}_u) = \sigma^2 x_u^T (X^T X)^{-1} x_u = \sigma^2 d^2(\hat{y}_u)$$

Si la fonction d'erreur de prédiction est plus petite que 1, on a une meilleure précision sur les réponses prédites que sur les réponses mesurées. Si la fonction d'erreur de prédiction est plus grande que 1, c'est l'inverse. En général, on s'arrange pour que l'erreur de prédiction ne soit pas plus élevée que l'erreur expérimentale. On évite donc de faire des prévisions dans les régions où la fonction d'erreur de prédiction est supérieure à l'unité.

3.2.2 Nombre de points marqués dans la configuration

Dans cette étude nous considérons une seule marque qui sera la valeur de l'erreur de prédiction, pour chaque point x_i on calcule la valeur de sa fonction d'erreur de prédiction.

En suite on calcule le nombre de points marqués dans la configuration de la manière suivante:

$$n(x) = \sum_{i=1}^n 1_{d(\hat{y}_{x_i}) \geq \varepsilon}$$

Avec ε une précision fixée. [16]

3.3 Simulation des processus ponctuels par la méthode MCMC et l'algorithme de Métropolies Hastings Green

Pour que le processus objet puisse atteindre la solution optimale, une solution classique consiste à définir une chaîne de Markov de type Monte Carlo (MCMC). Le but de ces méthodes est de construire une chaîne de configurations $\{X_0, X_1, \dots, X_N\}$ qui converge vers une distribution recherchée π . Pour cela en utilisant l'algorithme de Métropolies Hastings Green pour simuler une telle chaîne par construction d'un noyau de transition P qui est π -réversible, l'algorithme se décrit en 2 étapes :

- On fait une proposition de changement d'états x à y selon une loi de probabilité $Q(x, \cdot)$
- On accepte y avec la probabilité $a(x, y)$; sinon, on reste dans l'état x . Où $a: \Omega \times \Omega \rightarrow]0, 1]$.

La donnée simultanée de Q et a caractérise entièrement l'algorithme de MHG ou la chaîne de Markov sous-jacente. Notant $q(x, y)$ la densité de $Q(x, \cdot)$, la transition P de MHG s'écrit :

$$P(x, y) = a(x, y)q(x, y) + \left[\int_{\Omega} 1 - a(x, z)q(x, z) dz \right] \delta_x(y)$$

Avec $\delta_x(\cdot)$ et la masse du point en x , pour simplifier les calculs, on utilise la mesure de Dirac en x ($\delta_x(y) = 1$ si $x \in y$ et 0 sinon).[14]

Le choix de (Q, a) assurera la π -réversibilité de P si l'équation d'équilibre suivante est satisfaite :

$$\forall x, y \in \Omega : \pi(x) \times q(x, y) \times a(x, y) = \pi(y) \times q(y, x) \times a(y, x).$$

Le choix de la probabilité d'acceptation a est plus limité : il est dicté essentiellement par l'objectif de simuler (asymptotiquement) une loi de probabilité π donnée. C'est le cas du choix usuel, où:

$$a(x, y) = \pi(y) \times q(y, x) / \pi(x) \times q(x, y).$$

Si q est symétrique, alors:

$$a(x, y) = \pi(y) / \pi(x).$$

Dans ce contexte, un état $X_{(N)}$ de la chaîne construit par l'algorithme est un plan de n expériences (ensemble des positions des n points). La souplesse de l'algorithme réside dans le fait qu'il est aisé en général de simuler une transition d'état à état de la chaîne. L'algorithme construit une chaîne de Markov $(X_N)_{N \geq 0}$ qui peut explorer toutes les configurations de l'espace χ de noyau de transition :

$$K(x, y) = P^n(x, y)$$

Où $P^n(x, y)$ représente le noyau de transition associé au n transitions élémentaire à chaque itération N de l'algorithme.

Initialisation : choisir un plan d'expérience $x = \{x_1, \dots, x_i, \dots, x_n\}$ selon une loi de

probabilité donnée

Pour $j = 1$ à N

Pour $i = 1$ à n

- Choisir une expérience x_i au hasard
- Simuler y_j selon la loi q
- Accepter le changement $x_i = y_j$ avec la probabilité

$$a(x, y) = \min(1, \beta^{n(y)-n(x)} \exp^{u(x)-u(y)})$$

avec $y = \{x_1, \dots, y_j, \dots, x_n\}$ et $\beta > 0$

Fin

Fin

Algorithme 3.1 : Métropolis Hastings Green, simulation de processus ponctuel marqué de Gibbs.

Avec q est une densité de probabilité facile à simuler, par exemple une densité uniforme sur $[0,1]^d$ où bien une densité gaussienne tronquée. A ce niveau la question essentielle qui se pose est de savoir si la chaîne converge vers la distribution $\pi(x)$ définie en (1). (voir le paragraphe suivant) [15]

3.4 Etude théorique de la planification d'expériences par méthodes McMC

Pour assurer la convergence d'un processus il faut vérifier que la chaîne puisse atteindre la distribution invariante π à partir de n'importe quel état initial. donc il est nécessaire de pouvoir s'assurer de la convergence de la chaîne vers sa distribution finale. L'équation peut être écrite comme :

$$K^t(x, A) \xrightarrow[t \rightarrow \infty]{} \pi(A)$$

Où A un borélien de \mathcal{B} et $K^t(x, A) = K(X_t = A / X_0 = x)$ est un noyau de transition de pas t .

Enonçons le résultat principal qui nous intéresse ici :

Théorème 3.1

Si le noyau de transition $K = P^n$ est primitif sur un espace fini a une distribution unique invariante π alors la chaine de Markov $(X_k)_{k \geq 0}$ obtenue à partir de l'algorithme MH est uniformément ergodique et ce noyau réalise la simulation du PPM de Gibbs de densité $\pi(x) = k\beta^{n(x)} \exp^{-U(x)}$ (i.e. $\nu K^t \xrightarrow[t \rightarrow \infty]{} \pi$, pour toute distribution initiale ν)

Preuve

Premièrement nous montrons trois propriétés importantes pour le noyau P : π - réversibilité , π - stationarité et π - irréductibilité,

a. La π -réversibilité

La transition P est π -réversible si $\forall x, y \in \Omega : \pi(x)p(x, y) = \pi(y)p(y, x)$. Soit $x \in \Omega$ et $B \in \mathcal{B}$ on a :

$$\begin{aligned} \int_{\Omega} 1_{B(x,y)} \pi(x) P(x, y) dx &= \int_{\Omega} 1_{B(x,y)} \pi(x) a(x, y) q(x, y) dx + \int_{\Omega} 1_{B(x,y)} \pi(x) \left[\int_{\Omega} 1 - a(x, z) q(x, z) dz \right] \delta_x(y) dx \\ \int_{\Omega} 1_{B(x,y)} \pi(x) P(x, y) dx &= \int_{\Omega} 1_{B(x,y)} \pi(x) a(x, y) q(x, y) dx + \int_{\Omega} 1_{B(x,y)} \delta_x(y) \pi(x) \left[\int_{\Omega} 1 - a(x, z) q(x, z) dz \right] dx \\ \int_{\Omega} 1_{B(x,y)} \pi(x) P(x, y) dx &= \int_{\Omega} 1_{B(x,y)} \pi(x) a(x, y) q(x, y) dx + \int_{\Omega} 1_{B(x,x)} \pi(x) \left[\int_{\Omega} 1 - a(x, z) q(x, z) dz \right] dx \end{aligned}$$

Comme on a :

$$\pi(x) a(x, y) q(x, y) = k \beta^{n(x)} \exp^{-U(x)} \min\{1, \beta^{n(y)-n(x)} \exp^{U(y)-U(x)}\} q(x, y)$$

$$\pi(x) a(x, y) q(x, y) = k \min\{\beta^{n(x)} \exp^{-U(x)}, \beta^{n(y)} \exp^{-U(y)}\} q(x, y)$$

$$\pi(x) a(x, y) q(x, y) = k \beta^{n(y)} \exp^{-U(y)} \min\{\beta^{n(x)-n(y)} \exp^{U(y)-U(x)}, 1\} q(x, y)$$

En utilisant le faite que $q(x, y) = q(y, x)$, alors on aura $\pi(x) a(x, y) q(x, y) = \pi(y) a(y, x) q(y, x)$, d'où le résultat:

$$\begin{aligned} \int_{\Omega} 1_{B(x,y)} \pi(x) P(x, y) dx &= \int_{\Omega} 1_{B(x,y)} \pi(y) a(y, x) q(y, x) dx + \int_{\Omega} 1_{B(y,y)} \pi(y) \left[\int_{\Omega} 1 - a(y, z) q(y, z) dz \right] dy \\ \int_{\Omega} 1_{B(x,y)} \pi(x) P(x, y) dx &= \int_{\Omega} 1_{B(x,y)} \pi(y) P(y, x) dy \end{aligned}$$

Donc $\pi(x)p(x, y) = \pi(y)p(y, x)$, alors la chaine est π -réversible donc π -invariante.

b. La π -stationnarité

La transition P est π -stationnaire si $\pi P = \pi$, soit $x \in \Omega$ et $B \in \mathcal{B}$ on a :

$$\begin{aligned} \int_{\Omega} 1_{B(x,y)} \pi(x) P(x,y) dx &= \int_{\Omega} 1_{B(x,y)} \pi(x) a(x,y) q(x,y) dx + \int_{\Omega} 1_{B(x,y)} \pi(x) \left[\int_{\Omega} 1 - a(x,z) q(x,z) dz \right] \delta_x(y) dx \\ &= \int_{\Omega} \int_{\Omega} 1_{B(x,y)} \pi(x) a(x,y) q(x,y) dy dx + \int_{\Omega} 1_{B(x,x)} \pi(x) dx - \int_{\Omega} \int_{\Omega} a(x,z) q(x,z) \pi(x) dz dx \\ \int_{\Omega} 1_{B(x,y)} \pi(x) P(x,y) dx &= \int_{\Omega} 1_{B(x,x)} \pi(x) dx \end{aligned}$$

Donc la chaîne admet π comme distribution stationnaire.

c. La π -irréductibilité

La transition P est π -irréductible si

$$\forall A \in \mathcal{B} \text{ tq } \pi(A) > 0 \rightarrow \exists t \text{ tq } P^t(x, A) > 0$$

Soit $A \in \mathcal{B}$ et pour $t=1$ on obtient :

$$\begin{aligned} \int_{\Omega} 1_{B(x,A)} P(x, A) dx &= \int_{\Omega} 1_{B(x,A)} a(x, A) q(x, A) dx + \int_{\Omega} 1_{B(x,A)} \left[\int_{\Omega} 1 - a(x, z) q(x, z) dz \right] \delta_x(A) dx \\ \int_{\Omega} 1_{B(x,A)} P(x, A) dx &= \int_{\Omega} 1_{B(x,A)} a(x, A) q(x, A) dx + \int_{\Omega} 1_{B(x,x)} \left[\int_{\Omega} 1 - a(x, z) q(x, z) dz \right] dx \\ \int_{\Omega} 1_{B(x,A)} P(x, A) dx &= \int_{\Omega} 1_{B(x,A)} a(x, A) q(x, A) dx + 1 - \int_{\Omega} \int_{\Omega} a(x, z) q(x, z) dz dx \end{aligned}$$

On distingue quatre cas possibles :

- Si $a(x, A) = 1$ et $a(x, z) = 1$ alors

$$\int_{\Omega} 1_{B(x,A)} q(x, A) dx + 1 - \int_{\Omega} \int_{\Omega} q(x, z) dz dx$$

- Si $a(x, A) = 1$ et $a(x, z) = \gamma^{s(z)-s(x)} \beta^{n(z)-n(x)}$ alors

$$\int_{\Omega} 1_{B(x,A)} q(x, A) dx + 1 - \int_{\Omega} \int_{\Omega} \exp^{U(x)-U(z)} \beta^{n(z)-n(x)} q(x, z) dz dx$$

- Si $a(x, A) = \gamma^{s(A)-s(x)} \beta^{n(A)-n(x)}$ et $a(x, z) = 1$ alors

$$\int_{\Omega} 1_{B(x,A)} \exp^{U(x)-U(A)} \beta^{n(A)-n(x)} q(x, A) dx + 1 - \int_{\Omega} \int_{\Omega} q(x, z) dz dx$$

- Si $a(x, A) = \gamma^{s(A)-s(x)} \beta^{n(A)-n(x)}$ et $a(x, z) = \gamma^{s(z)-s(x)} \beta^{n(z)-n(x)}$ alors

$$\int_{\Omega} 1_{B(x,A)} \exp^{U(x)-U(A)} \beta^{n(A)-n(x)} q(x,A) dx + 1 - \int_{\Omega} \int_{\Omega} \exp^{U(x)-U(z)} \beta^{n(z)-n(x)} q(x,z) dz dx$$

$$\rightarrow \int_{\Omega} 1_{B(x,A)} P(x,A) dx = \begin{cases} \int_{\Omega} \int_A q(x,y) dx dy + 1 - \int_{\Omega} \int_{\Omega} q(x,z) dz dx \\ \int_{\Omega} \int_A q(x,y) dx dy + 1 - \int_{\Omega} \int_{\Omega} d_2 q(x,z) dz dx \\ \int_{\Omega} \int_A d_1 q(x,y) dx dy + 1 - \int_{\Omega} \int_{\Omega} q(x,z) dz dx \\ \int_{\Omega} \int_A d_1 q(x,y) dx dy + 1 - \int_{\Omega} \int_{\Omega} d_2 q(x,z) dz dx \end{cases}$$

$$\rightarrow \int_{\Omega} 1_{B(x,A)} P(x,A) dx = \begin{cases} \int_{\Omega} \int_A q(x,y) dx dy \\ \int_{\Omega} \int_A q(x,y) dx dy + 1 - d_2 \\ d_1 \int_{\Omega} \int_A q(x,y) dx dy \\ d_1 \int_{\Omega} \int_A q(x,y) dx dy + 1 - d_2 \end{cases}$$

$$\rightarrow \left[\int_{\Omega} 1_{B(x,A)} P(x,A) dx \right]^t = \begin{cases} \left[\int_{\Omega} \int_A q(x,y) dx dy \right]^t > 0 \\ \left[\int_{\Omega} \int_A q(x,y) dx dy + 1 - d_2 \right]^t > 0 \\ \left[d_1 \int_{\Omega} \int_A q(x,y) dx dy \right]^t > 0 \\ \left[d_1 \int_{\Omega} \int_A q(x,y) dx dy + 1 - d_2 \right]^t > 0 \end{cases}$$

Donc $\int_{\Omega} 1_{B(x,A)} P^t(x,A) dx > 0 \forall t \geq 0$, alors la chaîne est π -irréductible.

La transition P est π -irréductible et π -invariante. D'après le théorème 2.1 énoncé dans le chapitre 2 alors P est récurrente positive et π est l'unique loi invariante de P . L'irréductibilité de P entraîne la propriété de la primitivité. Donc pour $t \geq 1$, on a $K^t(x, \cdot) \geq 0$.

Si π est la loi invariante de P alors elle est loi invariante pour $K = P^n$. En effet, $\pi P = \pi$ (stationnarité de p) et par récurrence sur l'entier n , on obtient $\pi P^n = \pi$. Donc $\pi K = \pi$ (stationnarité de K)

Soit ν une distribution initiale, pour tout entier r nous avons :

$$\|\nu K^r - \pi\| = \|\nu K^r - \pi K^r\|$$

D'après le lemme 2.1 (chapitre 2) nous avons: $\|\nu K^r - \pi K^r\| \leq 2C(K^r)$ avec $C(K) = \frac{1}{2} \max_{x,y} \|K(x, \cdot) - K(y, \cdot)\|$ est le coefficient de contraction de Dobrushin 1956. Par récurrence sur r on obtient : $C(K^r) \leq C(K)^r$ donc :

$$\|\nu K^r - \pi\| \leq 2C(K^r) \leq 2C(K)^r$$

Comme $0 < C(K) < 1$, alors la chaîne est uniformément ergodique.

Pour r très grand, $\|\nu K^r - \pi\| \rightarrow 0$, π est l'unique loi invariante de K . La chaîne converge vers la distribution marquée de Gibbs.

CHAPITRE 4

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 des différentes méthodes utilisées dans la comparaison sont implémentés en MATLAB. Les programmes prennent comme donnée d'entrée des matrices d'expériences contenant des points générés d'une manière aléatoire, selon une densité donnée (en général on utilise la loi uniforme).

4.1 Matlab comme outil de programmation

Matlab est un langage hautes performances pour le calcul scientifique et technique. Il intègre la possibilité de calculs, de visualisation et de programmation dans un environnement très simple d'emploi. Les résultats sont exprimés sous une forme mathématique standard. L'utilisation typique est :

- Calcul scientifique.
- Développement d'algorithmes.
- Acquisition de données.
- Modélisation et simulation.
- Analyse de données, exploration et visualisation.
- Graphisme scientifique.
- Développement d'applications, interface graphique.

Matlab est un système interactif dont la brique de base est un tableau dont la taille n'est pas nécessairement connue. Ceci permet de résoudre des problèmes, en particulier ceux qui ont une formulation matricielle, en un minimum de temps (contrairement aux langages de bas niveau comme le C ou le fortran). Matlab a été développé à l'origine pour avoir un accès simple et rapide aux projets EISPACK et LINPACK. Aujourd'hui, Matlab intègre les bibliothèques LAPACK et BLAS incorporant ainsi les dernières techniques pour le calcul matriciel.

Dans l'enseignement universitaire Matlab s'est imposé comme un standard pour l'apprentissage de l'algorithmique scientifique. Dans l'industrie, il est l'outil de choix pour une productivité accrue en recherche et développement. Matlab peut aussi être enrichi à l'aide de Toolbox (boîtes à outils) pour des problèmes spécifiques. Matlab fournit une documentation détaillée à la fois en ligne et sous forme de fichiers PDF.

4.2 Etude de comparaison avec les plans usuels

Avant de comparer les plans en dimension supérieure où égale à 2, il est important de faire appel à des critères usuels permettant de juger le bon remplissage de l'espace ainsi la distribution uniforme. (Pour plus de détail voir le chapitre 1)

- Mesure de recouvrement: permet de mesurer l'écart entre les points du plan et ceux d'une grille régulière. Ce critère est nul pour une grille régulière. L'objectif est donc de le minimiser pour se rapprocher d'une grille régulière, et ainsi assurer le remplissage de l'espace, sans toutefois l'atteindre pour respecter une distribution uniforme notamment en projection sur les axes factoriels.
- Critère de distance: Le critère maximin (Mindist) consiste à maximiser la distance minimale entre deux points du plan.
- Critère de discrédance : La discrédance permet de mesurer l'écart entre la fonction de répartition empirique des points du plan et celle de la loi uniforme. Contrairement aux deux critères précédents, la discrédance n'est pas basée sur la distance entre les points. Il existe différentes mesures de discrédance. Nous retenons la discrédance en norme L2.
- Critère D-optimal: Une matrice d'expériences est dite *D-optimale*, maximiser le déterminant de $(X^T X)$ revient à minimiser tous les termes de $(X^T X)^{-1}$ et donc toutes les variances et covariances des coefficients.

Les plans proposés dans ce mémoire sont comparés avec les plans usuellement utilisés en expériences numériques:

- Plans aléatoires.
- Hyper cubes latins (LHS).
- Suite de Halton.
- Suite de Sobol.
- Plans de Strauss.
- Plans marqués de Gibbs.

4.3 Résultats et discussions

Dans cette partie, nous présentons les tables regroupant les résultats d'implémentation des différents algorithmes pour les différents plans considérés. Ces tables sont divisées selon le nombre de points et la dimension de l'espace expérimental. Les plans proposés donnent incontestablement les meilleurs résultats quel que soit le critère. Ils semblent concurrencer les plans de Strauss construit par l'approche utilisant des processus ponctuels simple. Même donnent un meilleur résultat que la suite de faible discrédance traditionnellement utilisés en phase exploratoire et ceci même pour le critère de discrédance.

4.3.1 Résultat pour 10 points et 2 dimensions

Tableau4.1 : listes des différents plans selon les quatre critères avec $n = 10$, $R = 0.1$, $\gamma = 0.01$, $\beta = 7$
 $NMC = 500$, $d = 2$

	Le rapport	discrédance	recouvrement	Distance
aléatoire	3,691143566	0,009749255	1,016573526	0,115767539
LHS	3,685942748	0,021284587	1,126300908	0,114261276
Halton	2,606358273	0,011660505	0,691795107	0,167244369
Sobol	3,115542728	0,00575314	1,436827018	0,088388348
Strauss	2,235834421	0,003321018	0,582012659	0,200467399
Gibbs marqué	<u>1,559603368</u>	<u>0,002627064</u>	<u>0,342718995</u>	<u>0,263540542</u>

Dix plans de chaque type sera générés afin de donner un sens aux résultats. Les figures ci-dessous représentent les box plots obtenus à la suite de cette étude pour les plans :

- Plans marqués de Gibbs.(1)

- Plans de Strauss.(2)
- Plans aléatoires.(3)
- Hyper cubes latins (LHS).(4)

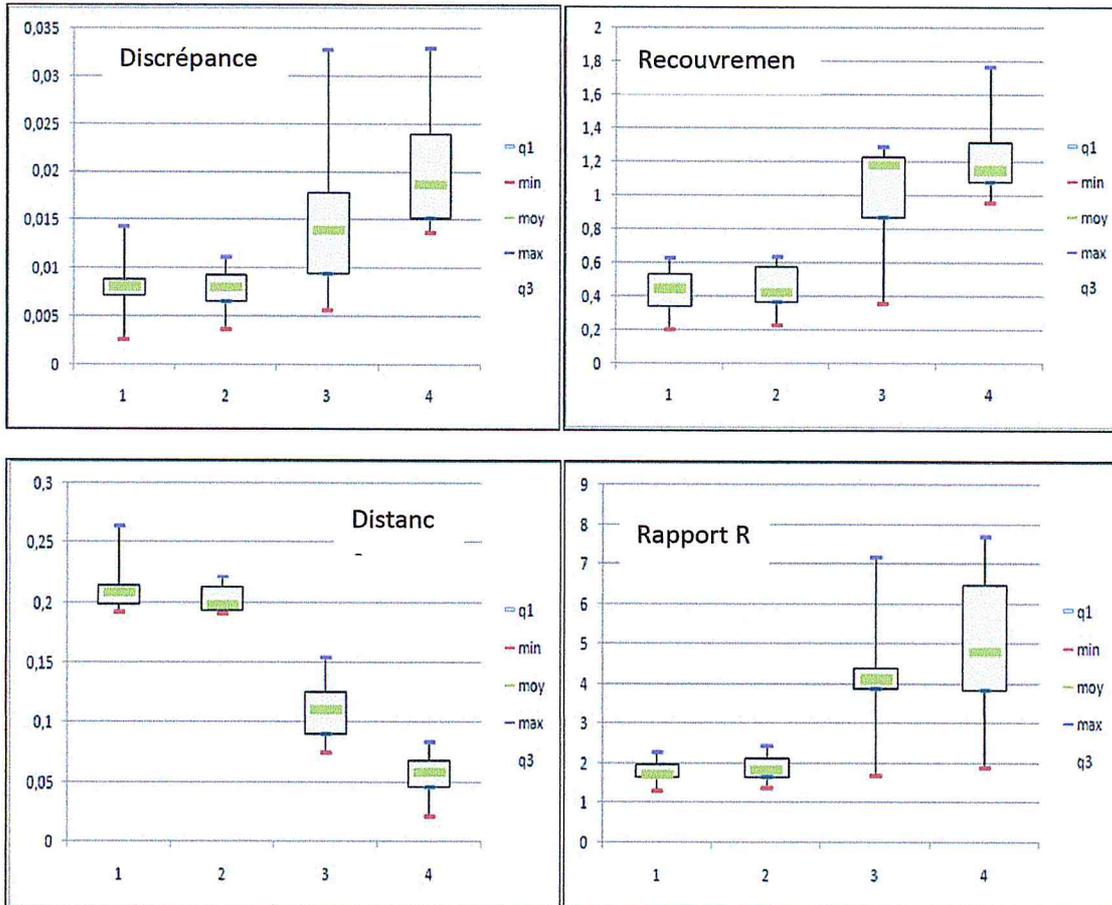


Figure 5.1 Box plots des critères de qualité pour deux dimensions

5.3.1 Résultat pour 20 points et 3 dimensions

Tableau 4.2 : listes des différents plans selon les quatre critères avec $n = 20$, $R = 0.1$, $\gamma = 0.01$, $\beta = 7$
 $NMC = 500$, $d = 3$

	Le rapport	discrépance	recouvrement	Distance
aléatoire	2,989969837	<u>0,003014416</u>	0,991834333	0,111513772
LHS	3,27918807	0,007502161	0,716467179	0,083774623
Halton	2,856179022	0,003894019	0,742467041	0,171961272
Sobol	2,311462122	0,003426396	0,702016899	0,184877493
Strauss	1,872518311	0,008437173	0,427444231	0,216928849
Gibbs marqué	<u>1,308478939</u>	0,007059617	<u>0,204033263</u>	<u>0,281193905</u>

Les figures ci-dessous représentent les résultats de 10 plans générés pour chaque types de plans suivants:

- Plans marqués de Gibbs.(1)
- Plans de Strauss.(2)
- Plans aléatoires.(3)
- Hyper cubes latins (LHS). (4)

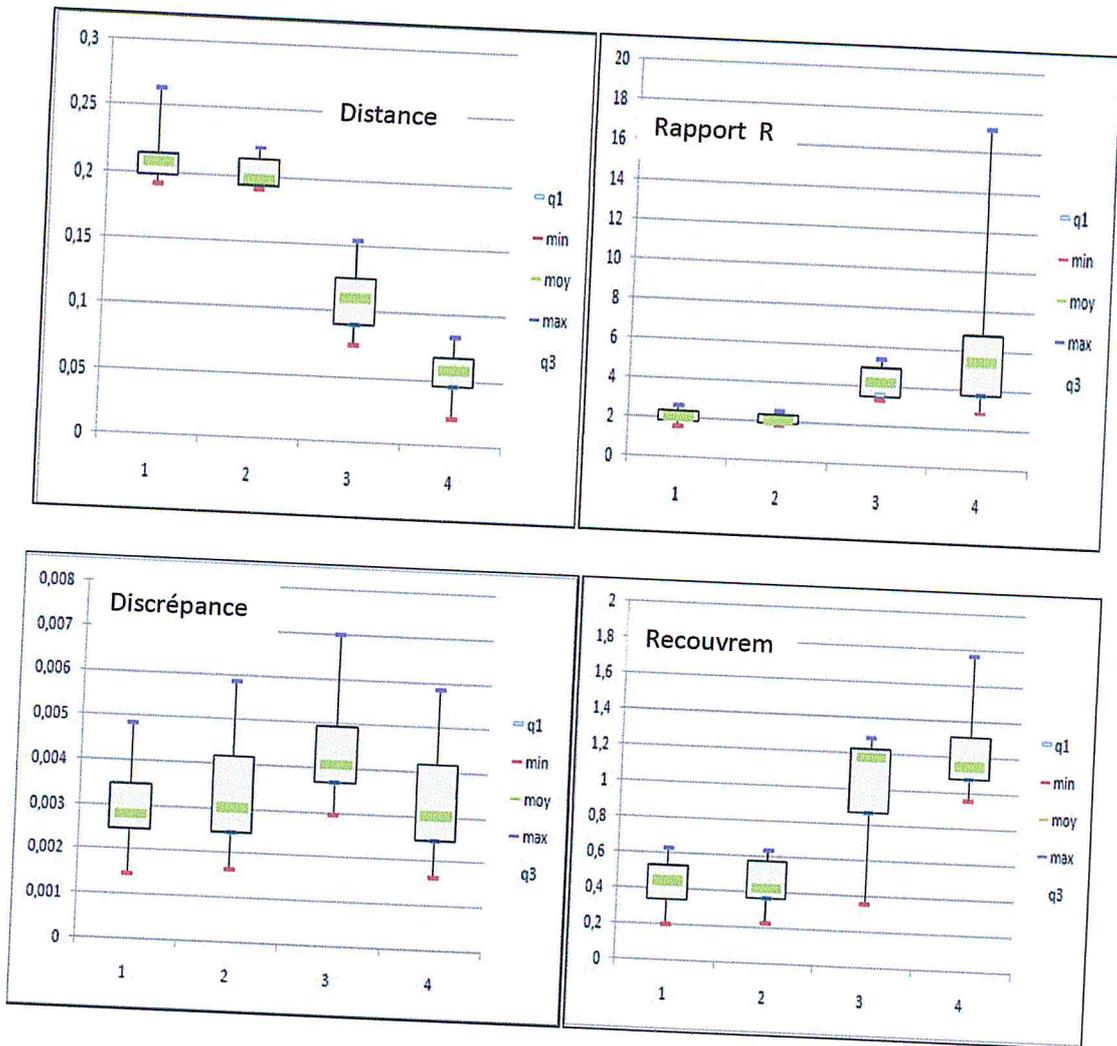


Figure 4.2 Box plot des critères de qualité pour trois dimensions

4.3.1 Résultat pour 50 points et 7 dimensions

Les résultats sont regroupés dans le tableau ci dessous, montrons une amélioration sur les valeurs des différents critères retenus.

Tableau 4.3 : listes des différents plans selon les quatre critères avec $n = 50$, $R = 0.19$, $\gamma = 0.05$, $\beta = 2.5$
 $NMC = 500$, $d = 7$

	Le rapport	discrépance	recouvrement	Distance
aléatoire	2,322480479	0,000129029	0,524327896	0,342672635
LHS	2,314952045	0,000269173	0,52709403	0,309870924
Halton	2,050373286	0,000631226	0,406518907	0,375855403
Sobol	2,178338009	0,000441577	0,414594945	<u>0,404744808</u>
Strauss	2,441723137	0,000117652	0,546702349	0,334557769
Gibbs marqué	<u>1,989474309</u>	<u>0,000120323</u>	<u>0,401501848</u>	0,365356689

Les figures ci-dessous représentent les résultats de 10 plans générés pour chaque types de plans suivants:

- Plans marqués de Gibbs.(1)
- Plans de Strauss.(2)
- Plans aléatoires.(3)
- Hyper cubes latins (LHS). (4)

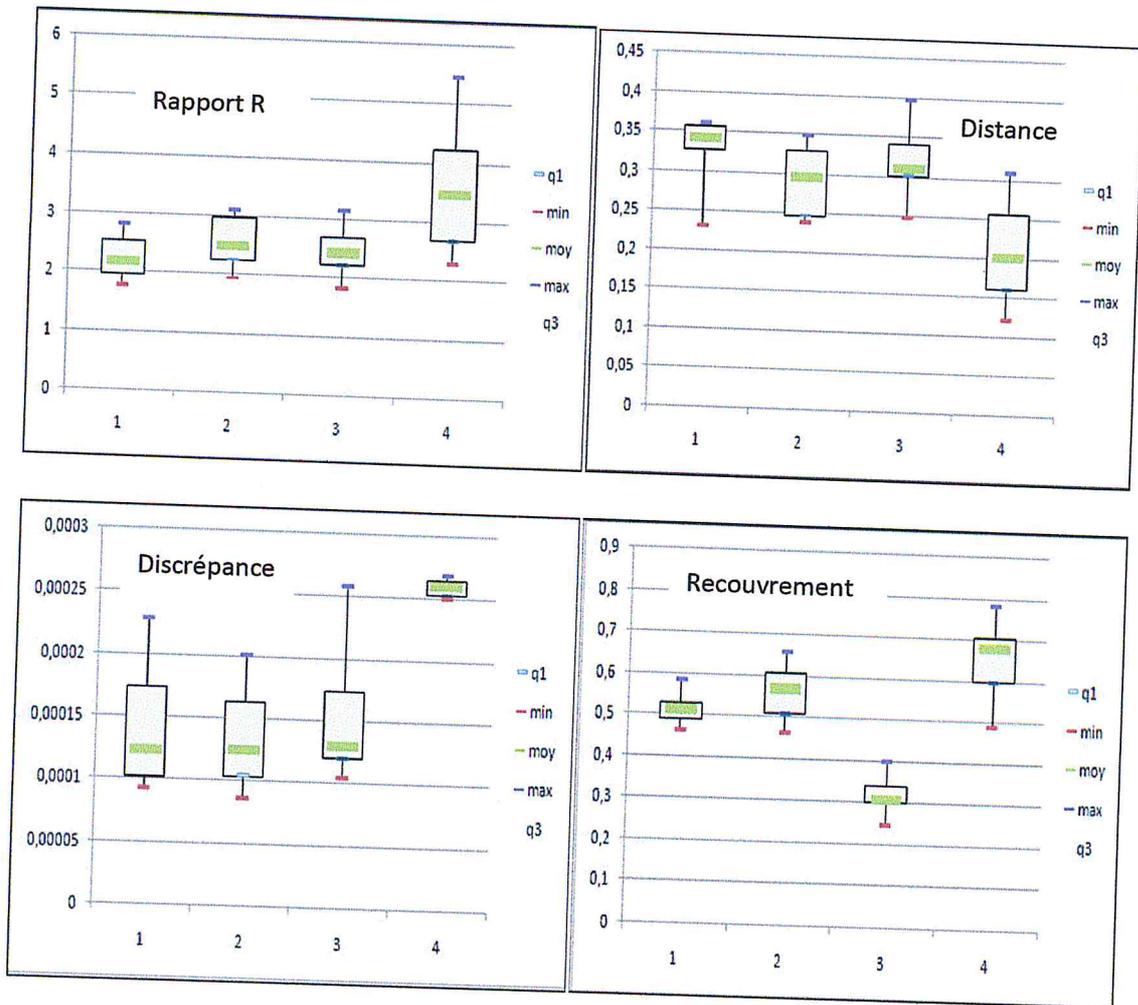


Figure 4.3 Box plot des critères de qualité pour sept dimensions

CONCLUSION GENERALE

En simulation numérique, les difficultés essentielles proviennent du coût du calcul du simulateur et de la grande dimension du problème à traiter. L'approche, actuellement privilégiée, consiste à construire un modèle d'approximation (régression, etc.) ou d'interpolation (krigeage, etc.) à partir d'un nombre réduit de simulations. Ces simulations sont effectuées dans des conditions fixées a priori et organisées selon un **plan d'expériences numériques**. Cette approche de modélisation de la réponse nécessite une connaissance a priori du type de dépendance entre la réponse et les paramètres et atteint ses limites lorsque la réponse étudiée est complexe. Ce travail se place dans le contexte de la phase exploratoire, à savoir lorsque cette dépendance n'est pas connue.

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 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).

Enfin, de nombreuses extensions pour la planification d'expériences numériques sont envisageables. Comme par exemple, l'utilisation de métriques différentes de la métrique euclidienne pour rendre compte d'hétérogénéité (notamment en grande dimension où les plans uniformes sont naturellement concentrés sur le bord du domaine).

REMERCIEMENT

Nous tenant à remercier sincèrement Monsieur, El Mossaoui Hichem, en tant que promoteur de mémoire, qu'est toujours montré à l'écoute et très disponible tout au long de la réalisation de ce mémoire, aussi pour l'inspiration, l'aide et le temps qu'il a bien voulu nous consacrer et sans qui ce mémoire n'aurait jamais vu le jour.

On n'oublie pas nos parents pour leur contribution, leur soutien et leur patience.

Enfin, nous adressons nos plus sincères remerciements à tous nos proches et amis, qui nous ont toujours encouragées au cours de la réalisation de ce mémoire.

Merci à tous et à toutes.

- [1] barbe P & ledoux M, probabilité, de la licence à l'agrégation, DUNOD paris 2004.
- [2] Barna Keresztes, segmentation de zones d'intérêt par processus objets, thèse doctorat à l'école doctorale des sciences physiques et de l'ingénieur 2009.
- [3] Bernard Bercu & Djalil Chafai, Modélisation Stochastique et Simulation, DUNOD, Paris 2007.
- [4] Carlo Gaetan & Xavier Guyon, Modélisation et Statistique Spatiales, Springer 2008.
- [5] Eric Thiémard, sur le calcul et la majoration de la discrédance à l'origine, thèse de doctorat à l'école polytechnique fédérale de Lausanne 2000.
- [6] Francois Jaulin, processus markoviens, thèse doctorat à l'école doctorale de centrale paris 2008.
- [7] Gerhard Winkler, Image Analysis Random fields and Dynamic Monte Carlo Methods 2000.
- [8] GOUPY Jacques. LES PLANS D.EXPERIENCES 2006.
- [9] Guillaume Perrin, étude du couvert forestier par processus ponctuels marqués, thèse doctorat à l'école doctorale de centrale paris 2006.
- [10] Jean-Jaques Droesbeke Michel Leijeune Gilbert Saporta, Analyse Statistique des Données Spatiales, TECHNIP, Paris 2006.
- [11] Jesper Moller & Rasmus Plenge Waagepetersen, Statistical Inference and Simulation for Spatial Point Processes, CHAPMAN & HALL/CRC 2003.
- [12] Jessica Franco, planification d'expérience numérique en phase exploratoire pour la simulation des phénomènes complexes, thèse de doctorat à l'Ecole Nationale Supérieure des Mines de Saint-Etienne 2009.
- [13] Krishna B Hani Doss, on the convergence of the Markov Chain simulation Method 1992.
- [14] Luke Tierney 1992, a note on Metropolis-Hastings kernels for general state spaces.
- [15] Luke tierney 1994, Markov Chauns for exploring posterior distribution.
- [16] Mikael Imberty & Xavier Descombes, simulation de processus objets.
- [17] Stéphane Gazut, Conception et mise en œuvre de nouvelles méthodes d'élaboration de plans d'expériences pour l'apprentissage de modèles non linéaires, thèse de doctorat à Sciences et Technologies de l'Information des Télécommunications et des Systèmes 2007.
- [18] Stéphane Vivier, stratégies d'optimales par plans d'expériences, thèse doctorat à l'école Doctorale Sciences et Ingénierie pour l'Information 2002.
- [19] Strauss DJ (1975) a model for clustering, Biometrika.

