

Détection et optimisation d'une zone vide

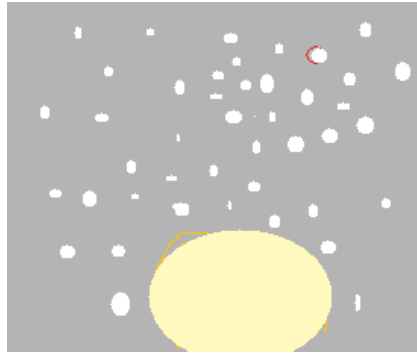
-Détermination de la répartition des points et recherche de zones moins denses



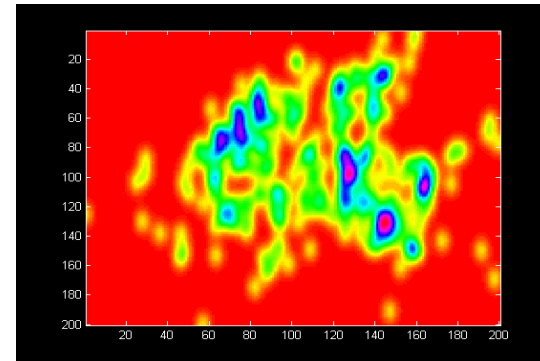
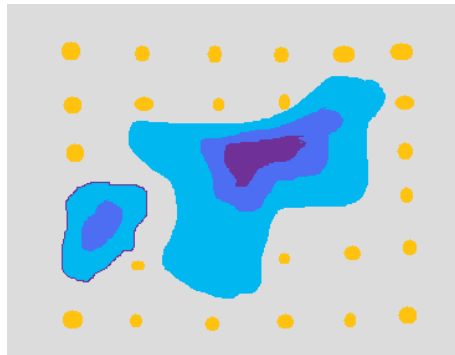
Quelques réflexions, pas plus

Au moins deux motivations :

-Analyser la structure, détection des zones vides, optimisation dans ces zones



-Ensemble de calibration : enrichissement, dé-densification, répartition uniforme



Détection et optimisation d'une zone vide

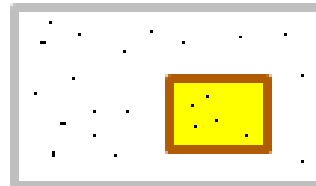
-Détermination de la répartition des points et recherche de zones moins denses

Définition:

-Discrépance :

« mesure entre l'uniformité parfaite et une configuration données (celle de la base de données)

« différence entre nombre de points compris dans un intervalle et le volume de cet intervalle »



Différentes définitions L^2 , modifiée, centrée, symétrique

« la discrétion est une mesure de la non-uniformité d'une séquence de points distribués dans un cube multidimensionnel »

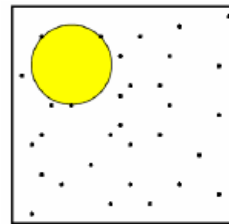
-Dispersion :

Permet d'évaluer les « trous », le rayon de la plus grande boule ne contenant aucune point de la base de données



extrait : Vincent Feuillard

« rayon de la plus grande boule ne contenant aucun point »



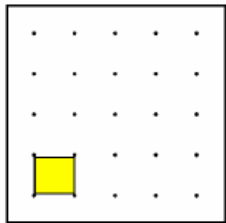
Définition : pour une distance d dans $[0,1]^d$, la dispersion de la suite

$$\{x_1, \dots, x_n\} \text{ est : } disp = \sup_{x \in [0,1]^d} \min_{x_i \in \{x_1, \dots, x_n\}} d(x_i, x)$$

Propriétés :

– une suite à discrédance faible est une suite à dispersion faible

– dans le cas de la distance d_∞ , il existe une borne minimale : $disp_\infty \geq \frac{1}{2[n^{1/d}]}$



Grille de Sukharev
($n=25$)

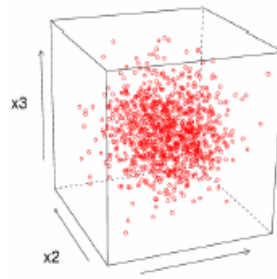
Ne pouvant calculer le supremum, nous approximos la dispersion par

$$dispR(x, x_f) = \max_{x_j \in \{x_{f_1}, \dots, x_{f_n}\}} \min_{x_i \in \{x_1, \dots, x_n\}} d(x_i, x_{f_j})$$

L'un des objectifs : évaluation de la structure et/ou de la qualité d'un jeu de données

Analyse :

Calcul des différentes **discrépances** de la base de données initiale BDD et comparaison avec des suites à discrécances faible comportant le même nombre de points



	$DiscL^2$	$DiscL^2M$	$DiscL^2C$	$DiscL^2S$
Base de données	0.09564124	0.3330283	0.2562327	0.8243496
Suite de Halton	0.004167211	0.01093556	0.008276079	0.02737285
Suite de Hammersley	0.002862335	0.007592062	0.006524325	0.02121937
Suite de Faure	0.003910729	0.008620693	0.007300586	0.03005841
Réseau	0.003353597	0.006752704	0.006019821	0.02675386

Outil 1

x1 Utilisation de « ratio » pour comparer les discrécances de la base de données et celles des suites à discrécance faible.

Utilisation de ratio :

Les ratio que l'on utilise sont de la forme :

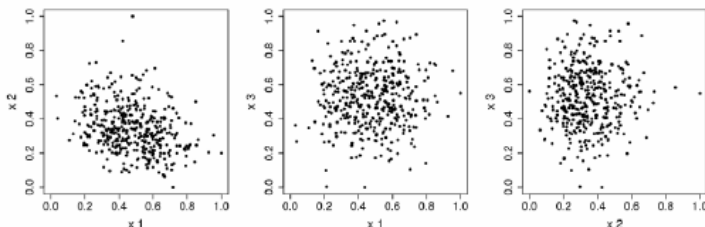
$$RDiscL^2_{min} = \frac{DiscL^2(BDD_{ini})}{\max_{\{Hal.,Ham.,Fa.,Res.\}} DiscL^2(SDF)} \quad \text{et} \quad RDiscL^2_{max} = \frac{DiscL^2(BDD_{ini})}{\min_{\{Hal.,Ham.,Fa.,Res.\}} DiscL^2(SDF)}$$

	$RDiscL^2$	$RDiscL^2M$	$RDiscL^2C$	$RDiscL^2S$
Minimum	22.95090	30.45370	30.96064	27.42492
Maximum	33.41371	49.31777	42.56485	38.84891

Ordre de grandeur important (/ suite à discrécance faible)

⇒ Il existe un intervalle où la densité de points est « anormale » (faible ou élevée) / au volume de cet intervalle

⇒ Recouvrement de l'espace peu satisfaisant



L'un des objectifs : évaluation de la structure et/ou de la qualité d'un jeu de données

Analyse : Outil 2

Calcul de la **distance minimale** entre deux points.

$$d_{min} = \min_{x_i, x_j \in \{x_1 \dots x_n\}} d(x_i, x_j)$$

Dans cet exemple: $d_{min}_2 = 0.009054198$ (distance euclidienne)

$$d_{min}_\infty = 0.007155492 \text{ (distance maximale)}$$

La valeur obtenue permet de savoir s'il y a redondance d'information (localement).

→ Nécessite l'avis d'expert (comportement de la physique)

→ sera utile pour apprécier la valeur de la dispersion

Autres distances possible : $\frac{2}{n(n-1)} \times \sum_{x_i, x_j \in \{x_1, \dots, x_n\}} \frac{s^{1/2}}{d_2(x_i, x_j)}$

(inversement proportionnel à la moyenne harmonique)

Outil 3

Calcul de la **dispersion « relative »** à une suite à discrédance faible :

$$dispR(x, x_f) = \max_{x_f \in \{x_1 \dots x_n\}} \min_{x \in \{x_1 \dots x_n\}} d(x_i, x_f)$$

ou x_f est une suite à discrédance faible (suite de Halton, Hammersley, Faure, Réseau ...)

DispHa	DispHam	DispFa	DispRes	$1/(2n^{1.4})$
0.4093491	0.3941218	0.3828636	0.3773609	0.0714287

Norme maximale

DispHa	DispHam	DispFa	DispRes
0.526165	0.5356336	0.556336	0.5305011

Norme euclidienne

Pour apprécier la dispersion

$$d_{min}_2 = 0.009054198 \text{ (distance euclidienne)}$$

$$d_{min}_\infty = 0.007155492 \text{ (distance maximale)}$$

⇒ des « trous »

Le rayon de la plus grande boule vide « bien » supérieur à la distance minimale entre deux points
⇒ De la redondance, densité importante de points dans certains intervalles

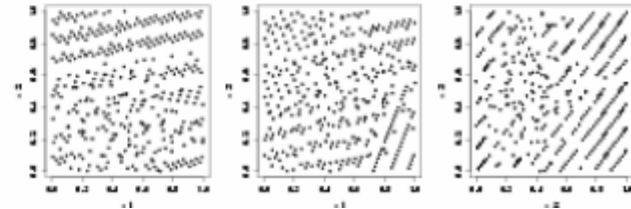
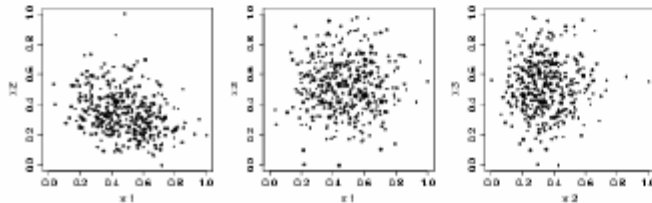
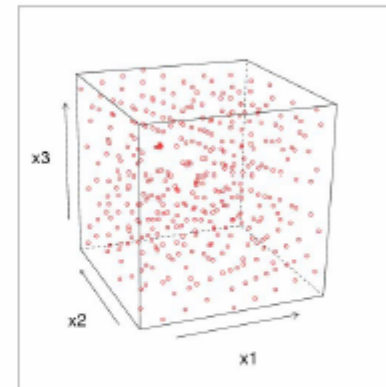
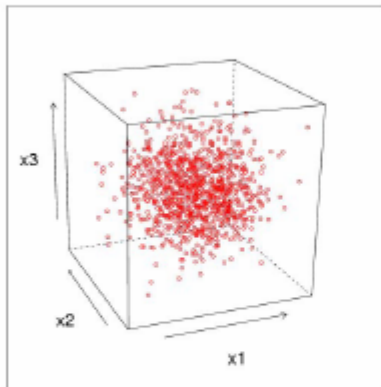
Extrait article
V Feuillard - CEA

Objectif final : aménager les zones vides, moins de points, mieux équirépartis

Comparaison visuelle



Moins de points mais mieux « équirépartis »



Autre bibliographie : (1)

Thèse d'Eric Thiémard (2000) - Ecole polytechnique de Lausanne
« sur le calcul et la majoration de la discrédance à l'origine »

-Rappel théorie de l'équirépartition, étude du concept d'uniformité

- la méthode de Monte-Carlo (*) et sa version déterministe (suites à discrédance faible : van der Corput, Halton, Hammersley, Sobol, Faure, ...)

-présentation de méthodes de calcul de discrédance, partition de cubes, géométrie combinatoire

- Approche par programmation linéaire en nombres entiers

Monte-Carlo (*) le principe de base consiste en l'utilisation du hasard pour aborder l'étude d'un problème déterministe

Autre bibliographie : (2)

STATISTIQUE SPATIALE

Xavier Guyon

SAMOS - Université Paris 1 Février 2007

Conférence S.A.D.A.' 07 - Cotonou - Benin

1 Géostatistique	2 Auto-régression et corrélation spatiales
1.1 Quelques rappels	2.1 Auto-régression stationnaire
1.2 Processus stationnaire, covariance	2.1.1 Moyenne glissante (MA) et ARMA
1.2.1 Définitions, exemples	2.1.2 Auto-régression simultanées (SAR)
1.2.2 Représentation spectrale d'une covariance	2.1.3 Auto-régression conditionnelle (CAR)
1.3 Processus intrinsèque, variogramme	2.2 Modèle non-stationnaire sur S fini
1.3.1 Exemples de variogramme	2.2.1 Modèle SAR
1.3.2 Propriétés géométriques : continuité, différentiabilité	2.2.2 Représentation markovienne CAR
1.4 Prédiction à covariance connue : le krigeage	2.2.3 Champ gaussien markovien
1.4.1 Le krigeage simple	2.2.4 AR en présence d'exogènes : modèle SARX
1.4.2 Le krigeage universel	2.3 Auto-corrélation spatiale latticielle
1.5 Estimation en géostatistique	2.3.1 L'indice de Moran
1.5.1 Analyse variographique	2.3.2 Test asymptotique d'indépendance spatiale
1.5.2 Estimation empirique du variogramme	2.3.3 L'indice de Geary
1.5.3 Validation d'un modèle de variogramme	2.3.4 Test de permutation d'indépendance spatiale
1.6 Exercices	2.4 Statistique d'un champ du second ordre
	2.4.1 Estimation d'un modèle stationnaire sur \mathbb{Z}^d
	2.4.2 Estimation d'un modèle auto-régressif
	2.4.3 Maximum par vraisemblance pour un champ gaussien centré
	2.4.4 Estimation d'une régression spatiale
3 Champ de Markov sur réseau	4 Processus ponctuels spatiaux
3.1 Recollement de lois conditionnelles	4.1 Définitions et notations
3.2 Champ de Gibbs sur S	4.1.1 Espace exponentiel des configurations
3.2.1 Potentiel d'interaction et spécification de Gibbs	4.1.2 Loi d'un processus ponctuel
3.2.2 Exemples de spécification de Gibbs	4.2 Exemples de processus ponctuels
3.3 Champ de Markov et champ de Gibbs	4.2.1 Processus Binomial
3.3.1 Cliques d'un graphe, champ de Markov	4.2.2 Processus Ponctuel de Poisson (PPP)
3.3.2 Le théorème de Hammersley-Clifford	4.3 Densité d'un processus ponctuel
3.4 Les auto-modèles de Besag	4.3.1 Définitions
3.4.1 Recollement de lois conditionnelles et auto-modèle	4.3.2 Exemples
3.4.2 Exemples d'auto-modèles	4.4 Caractéristiques du second ordre d'un PP
3.5 Estimation d'un champ de Markov	4.4.1 Moment et intensité d'ordre p
3.5.1 Le maximum de vraisemblance	4.4.2 Mesure de Palm
3.5.2 Pseudo-vraisemblance conditionnelle (PVC) de Besag	4.4.3 Deux distances au plus proche voisin
3.5.3 La méthode de codage	4.4.4 Moment réduit d'ordre 2 de Ripley
3.5.4 Précisions comparées du MV, MPVC et du Codage	4.5 Processus ponctuel de Markov
	4.6 Statistique des processus ponctuels spatiaux
	4.6.1 Estimation de l'intensité d'un PP
	4.6.2 Estimation de caractéristiques au second ordre
	4.6.3 Estimation par MC d'un modèle paramétrique
	4.6.4 Test de Monte Carlo
	4.6.5 Pseudo-vraisemblance d'un PP
	4.7 Calcul de Monte Carlo d'une vraisemblance

Autre bibliographie : (3)

Thèse de Jessica Franco (thèse Total EP)



- Planification d'expériences numériques à partir du processus de Strauss
- Analyse de la qualité intrinsèque des plans d'expériences numériques
« space filling design » de grande dimension par l'arbre de longueur minimal



THESE
présentée par

Jessica FRANCO

Pour obtenir le grade de DOCTEUR
de l'Ecole Nationale Supérieure des Mines de Saint-Etienne

Spécialité : *Mathématiques Appliquées*

*Planification d'expériences numériques
en phase exploratoire
pour la simulation des phénomènes complexes*

Préparée au département Méthodes et Modèles Mathématiques pour l'Industrie (3MI) de
l'ENSM-SÉ et au département Géomodélisation et Incertitudes (G&I) de TOTAL

Soutenue publiquement à Saint-Etienne le 10 septembre 2008

Exemples



Deux questions :

- 1- analyse de la structure
- 2- déterminer les zones vides

Deux autres questions :

- Travailler dans l'espace à p dimensions
- Ou travailler dans un espace à 2 dimensions (après projection) : *choix ici*

1- La complexité du calcul de la discrédance est dépendante de la Dimension et devient donc inenvisageable en grande dimension.

2- Comment remplir l'espace ?

Si l'on avait seulement deux variables :

-Plan factoriels complets, fractionnaires, Composites, Box-Behnken

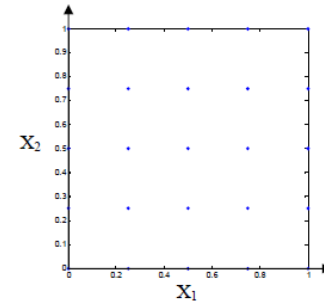
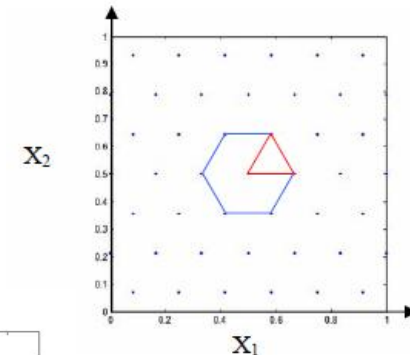
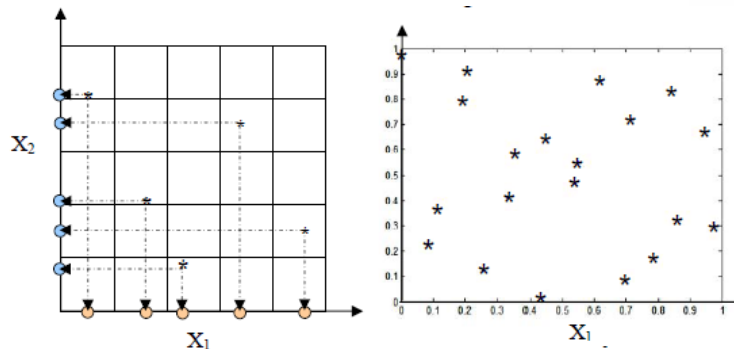


Figure 2. Un plan factoriel complet à 5 niveaux.

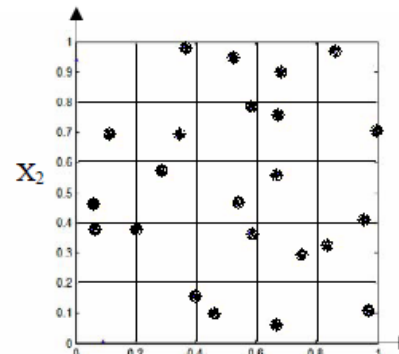
-Réseaux uniformes de Doehlert



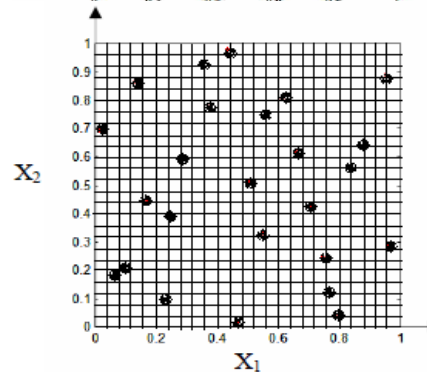
-Hypercube latins



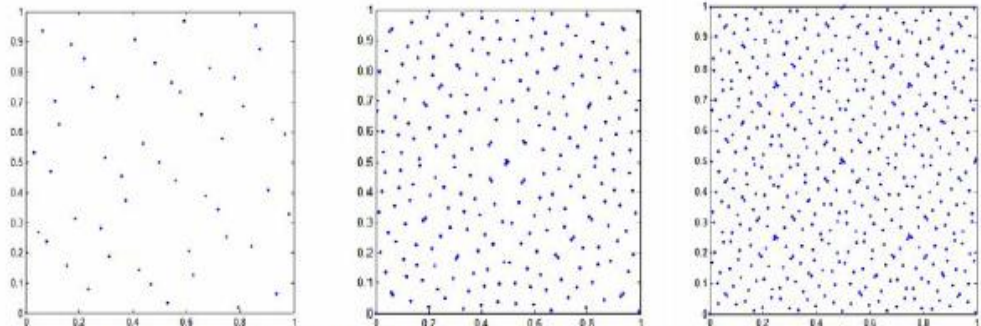
-Tableaux orthogonaux



-Hypercubes latins basés sur les tableaux orthogonaux



Suites à faible **discrepance** générées par des algorithmes déterministes afin d'obtenir des points répartis le plus uniformément possible dans le domaine expérimental.



Les 50, 250 et 500 premiers points d'une suite de Niederreiter en dimension 2

Etudier l'uniformité d'une distribution de points est un travail fastidieux, voire même impossible en grande dimension.

Critères généralement étudiés :

- Critères de discrédance qui permettent de mesurer si la distribution donnée est proche d'une distribution uniforme
- Classe des critères de distance qui comparent une distribution donnée à une grille régulière
- Critère d'entropie qui mesure la quantité d'information contenue dans le plan

En réalité la discrédance seule ne suffit pas à mesurer la qualité d'un Plan car on veut à la fois mesurer l'uniformité des points et mesurer le remplissage de l'espace.

Les plans WSP (Wooton Sergent Phan-Tan-Luu) qui repose sur l'algorithme de Kennard et Stone et Wooton

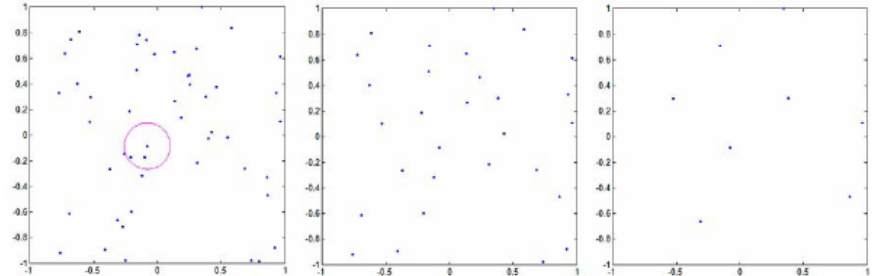


Figure 38. A gauche, un plan initial (aléatoire) de 50 points avec la première boule d'élimination. Au centre, le plan à 28 points obtenu avec la boule d'élimination de la figure de gauche. A droite, le plan à 8 points obtenu avec une boule d'élimination de rayon supérieur

Et à entropie maximale (mais présuppose un modèle sous-jacent)

Permettent de bien remplir l'espace mais méthodes liées au choix de la distribution initiale.

⇒ Méthodes usuelles non adaptées

⇒ méthode proposée : processus ponctuel de Stauss, plans construits à partir de méthodes Monte-Carlo par chaînes de Markov

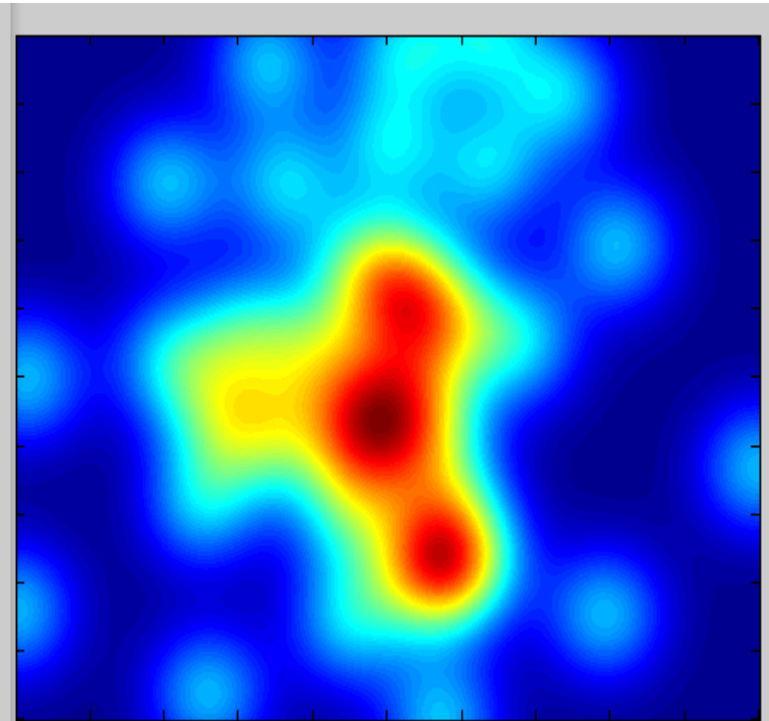
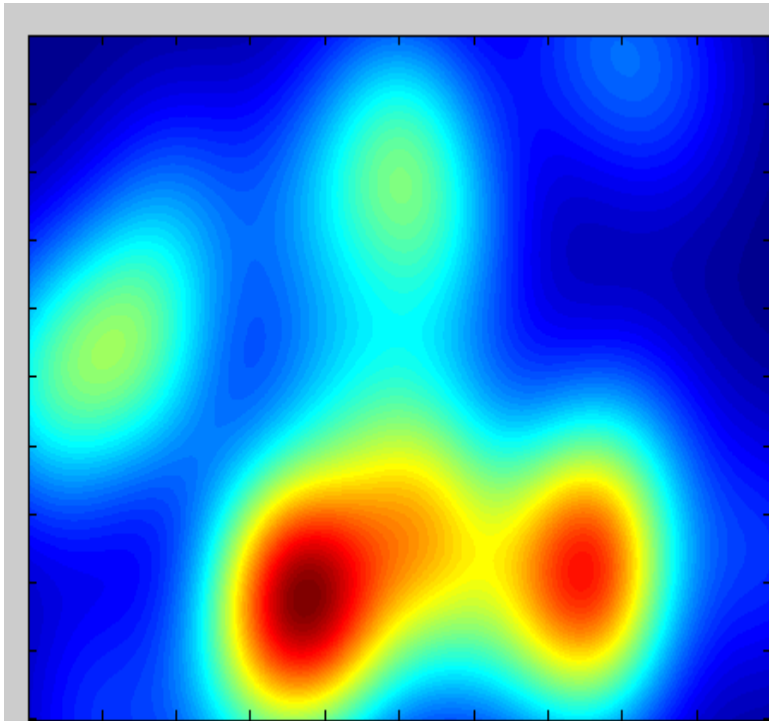
Concrètement ...

Commençons par l'analyse de la structure

Fonction Potentiel à partir de :

ACP (2 premières composantes)

ACC (méthode de projection
Non linéaire)

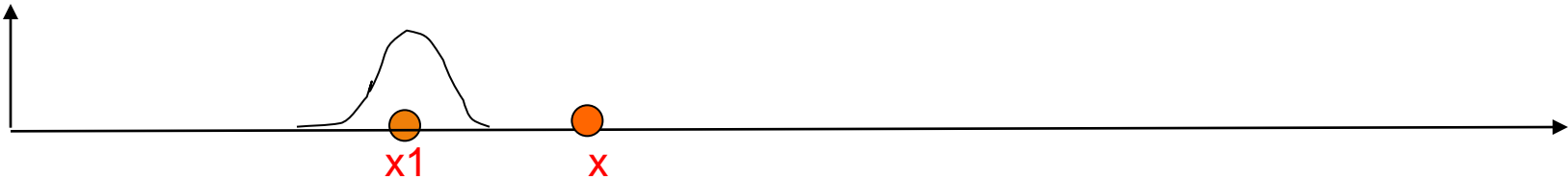


La démarche

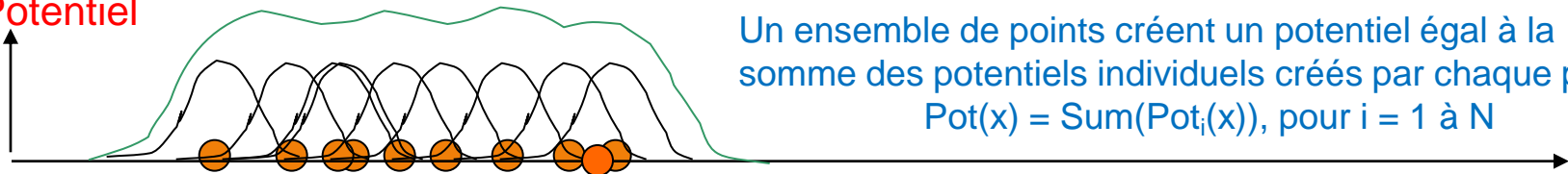
Expression générale du potentiel

$$Pot(x_i, x_j, u_i) = \prod_{m=1}^M \left(\frac{1}{u_i(m) \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{(X(i,m) - X(j,m))^2}{2u_i(m)^2}\right) \right)$$

Potentiel



Potentiel

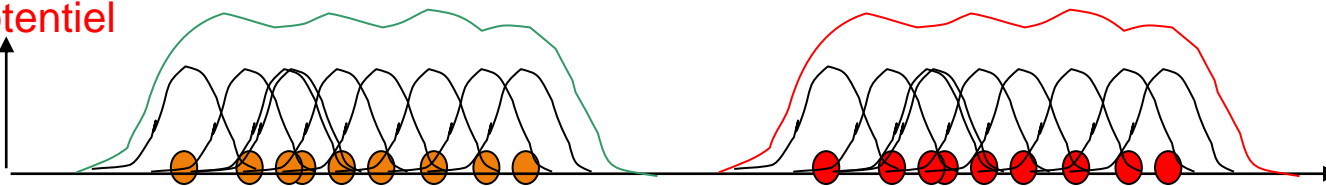


Un ensemble de points crée un potentiel égal à la somme des potentiels individuels créés par chaque point :
 $Pot(x) = \text{Sum}(Pot_i(x))$, pour $i = 1$ à N

Enveloppe des échantillons verts

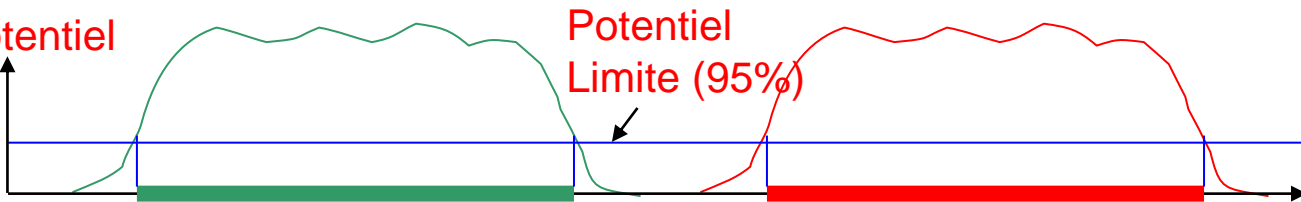
Enveloppe des échantillons rouges

Potentiel



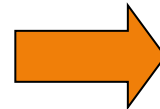
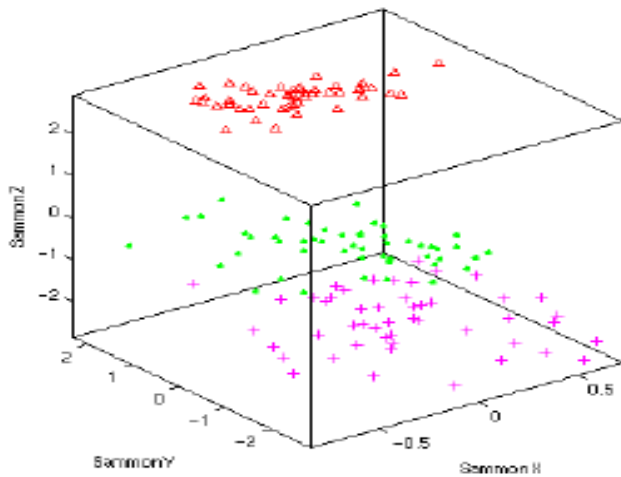
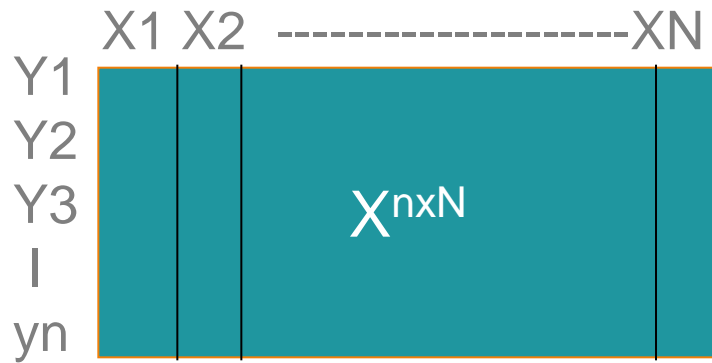
Pour deux groupes distincts, on calcule le potentiel créé par les points de chacun des groupes. La zone où le potentiel créé par les points rouges (resp verts) est élevé est caractéristique des points rouges (resp verts).

Potentiel

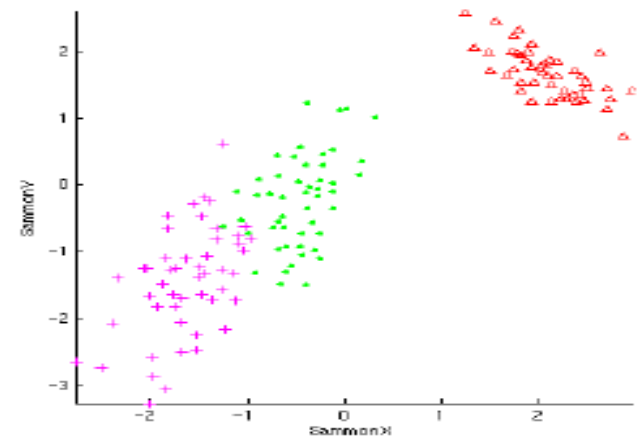


Domaine des échantillons verts

Domaine des échantillons rouges



Projection non linéaire



Méthodes essentiellement à base d'apprentissage garantissant la distance des points et les distorsions de données à structure Non-linéaire : ces méthodes permettent de visualiser des données à fortes dimensions et de reproduire toute l'information sur une carte en 2D

MDS : multidimensional scaling

CCA : analyse en composantes curvilignes

CDA : curvilinear distance analysis

LVQ : learning vector quantization

Carte de **Sammon**

Carte de **Kohonen**

Carte **GTM** : generative topographic mapping

Carte **RPM** : relational perspective map

LLE (méthode algébrique)

ISOMAP

HLLE (méthode algébrique)

les méthodes non-linéaires ont comme avantage de pouvoir, permettre de projeter efficacement des distributions plus complexes de données, au prix des difficultés généralement associées aux méthodes d'analyse non-linéaire (convergence, minima locaux, etc.). De plus, elles utilisent de plus en plus des notions de distances non-Euclidiennes.

Ces méthodes de projection non-linéaire sont basées sur des critères de conservation de la variance, des distances entre points, ou des voisinages (Multi-Dimensional Scaling, Curvilinear Component Analysis, Curvilinear Distance Analysis, Shannon mapping, Isomap, etc.).

Ces méthodes ont chacune leurs avantages et inconvénients.

L'analyse en composantes curvilignes (ACC)

Principe :

L'algorithme ACC par rapport aux cartes de Kohonen vise :

- A séparer les opérations de quantification et de projection
- A autoriser un choix de dimension de l'espace de projection
- A laisser la forme de la carte libre d'épouser au mieux la disposition des données

Le problème est de représenter le mieux possible un nuage de n_{obs} vecteurs d'observation de dimension n dans un espace de dimension inférieure p .

Pour ce faire, on dispose de N neurones qui n'ont pas de topologie définie a priori, celle-ci sera trouvée par apprentissage. Chaque neurone i possède un vecteur d'entrée x_i de dimension n qui pointe dans l'espace des données et un vecteur de sortie y_i de dimension $p < n$ qui pointe dans l'espace de représentation.

On procède d'abord à une quantification vectorielle de l'espace des données par les N vecteurs prototypes x_i soit selon l'un des algorithmes classiques, soit en imposant une répartition uniforme des prototypes. Ces vecteurs sont attribués au hasard aux neurones.

La quantification vectorielle (VQ) permet d'obtenir de petits ensembles (appelés centroides) permettant de gagner du temps sur le temps obtenu lors de la descente du gradient.

On note X_{ij} la distance euclidienne entre les deux vecteurs x_i et x_j .

Dans un second temps, on cherche à reproduire la topologie d'entrée dans l'espace de sortie en adaptant les vecteurs y_i de sortie de façon à ce que la distance Y_{ij} entre les prototypes y_i et y_j en sortie soit égale à la distance X_{ij} entre les deux prototypes en entrée x_i et x_j .

Si les dimensions de l'espace d'entrée et de l'espace de sortie sont identiques, cela revient à minimiser la forme quadratique suivante :

$$\sum_i \sum_{j \neq i} (X_{ij} - Y_{ij})^2$$

Dans le cas où ces dimensions sont différentes ($p < n$) on pourra pondérer ces termes pour favoriser les distances les plus faibles dans l'espace de représentation permettant un respect local de la topologie (zone très pliée en entrée).

$$E = \sum_i \sum_{j \neq i} (X_{ij} - Y_{ij})^2 F_\lambda(Y_{ij})$$

Les quantités $(X_{ij} - Y_{ij})^2$ vont donc être pondérées par un coefficient positif $F_\lambda(Y_{ij})$ qui favorise les distances courtes au dépend des longues distances. Cette fonction est appliquée sur les sorties car en principe, il est impossible de connaître la nature du plissement des données a priori.

L'idée est de comparer les distances de paires de points dans l'espace d'entrée dx et dans l'espace de sortie dy .

$F_\lambda(Y_{ij})$ est une fonction décroissante et monotone, son étendue est contrôlée par un paramètre λ ajustable.

La minimisation de E se fait au moyen de l'algorithme du gradient sur les composantes de chaque vecteur y_i (la minimisation de la fonction d'énergie par une méthode de type recuit simulé ne donne pas de bons résultats) :

En considérant que E est la somme d'énergies partielles :

$$E = \sum E_i$$

Le gradient de E par rapport au vecteur y_i s'exprime par une somme de gradients partiels.

L'algorithme peut se résumer de la manière suivante :

A chaque itération de l'algorithme :

- 1- un i ($1 \leq i \leq n_{obs}$) est fixé aléatoirement
- 2- tous les vecteurs y_j sont déplacés selon la règle :

$$\Delta y_j = \mu F_\lambda(Y_{ij})(X_{ij} - Y_{ij}) \frac{y_j - y_i}{Y_{ij}}$$

où

- M est un gain d'adaptation
- Le terme $F_\lambda(Y_{ij})(X_{ij} - Y_{ij})$ provient des dérivées partielles par rapport aux y_j
- $\frac{y_j - y_i}{Y_{ij}}$ est un déplacement unitaire

Qualité de représentation topologique :

L'idée est de comparer les distances de paires de points dans l'espace d'entrée dx et dans l'espace de sortie dy . On trace donc sur un graphique la distribution conjointe des distances pour tous les couples.

En résumé :

ACC demande 3 paramètres :

- la dimension de l'espace d'entrée p
- le facteur d'apprentissage $\alpha(t)$ (décroissance de 0.95 à 0.01)
- le facteur de voisinage $\lambda(t)$

Le dernier est plus critique car une descente trop rapide compromet la convergence ; le déploiement d'une spirale par exemple s'avère difficile par ACC.

Par contre une descente trop lente pénalise les grandes distances.



Méthodes

ACP modèle linéaire – utilise un critère basé sur la conservation de la variance

MDS utilise la similarité, les distances – méthode itérative – minimisation d'une fonction Stress

Isomap non prédictif – diffère de MDS par une métrique non euclidienne – le changement d'espace (initial/final) s'effectue de manière algébrique – sélection aléatoire des points candidats

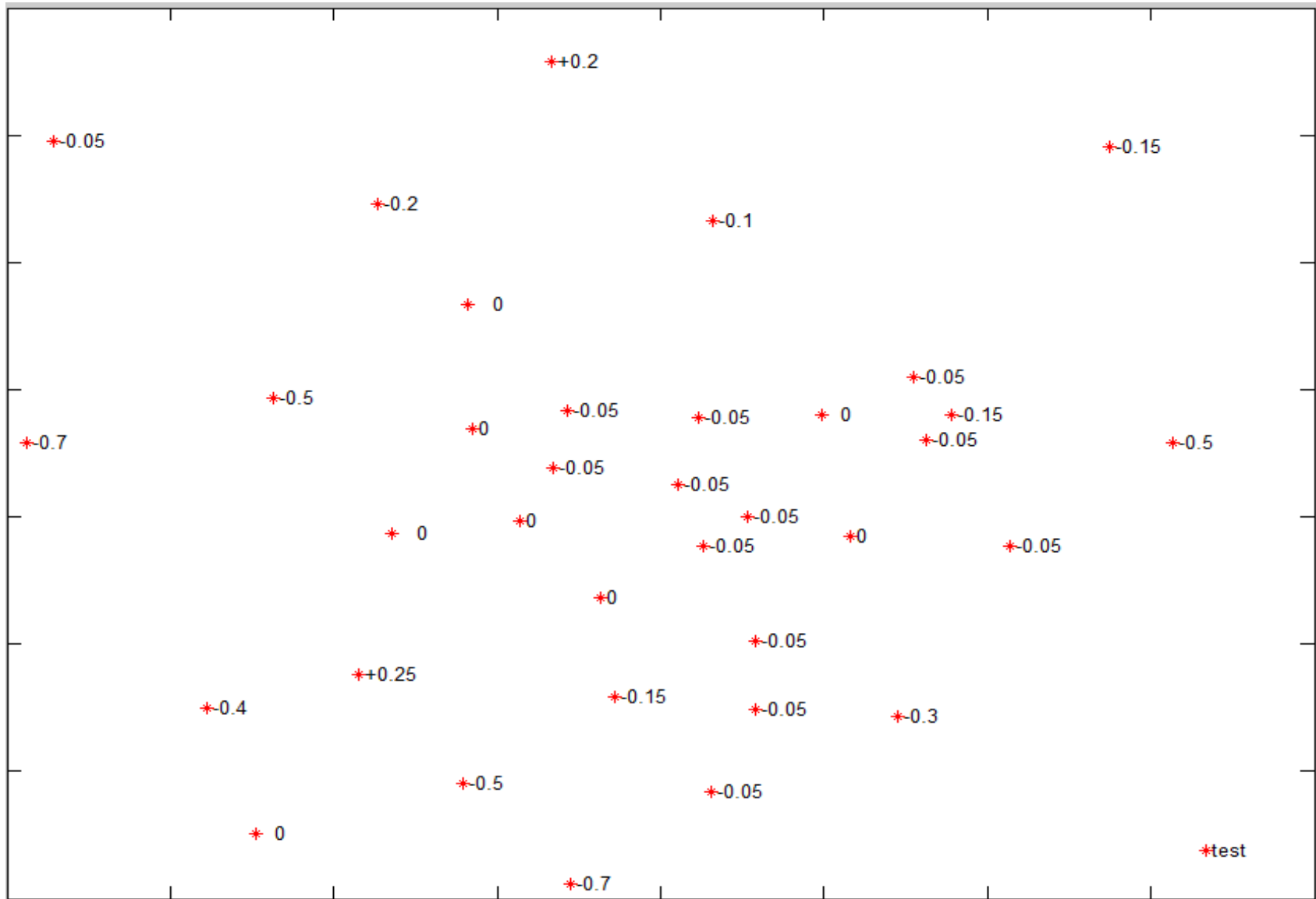
Sammon transformation non linéaire – algorithme par apprentissage – pas de projection d'un point nouveau sur la carte

CDA réseau de neurones – points candidats par quantification vectorielle – métrique par descente du gradient

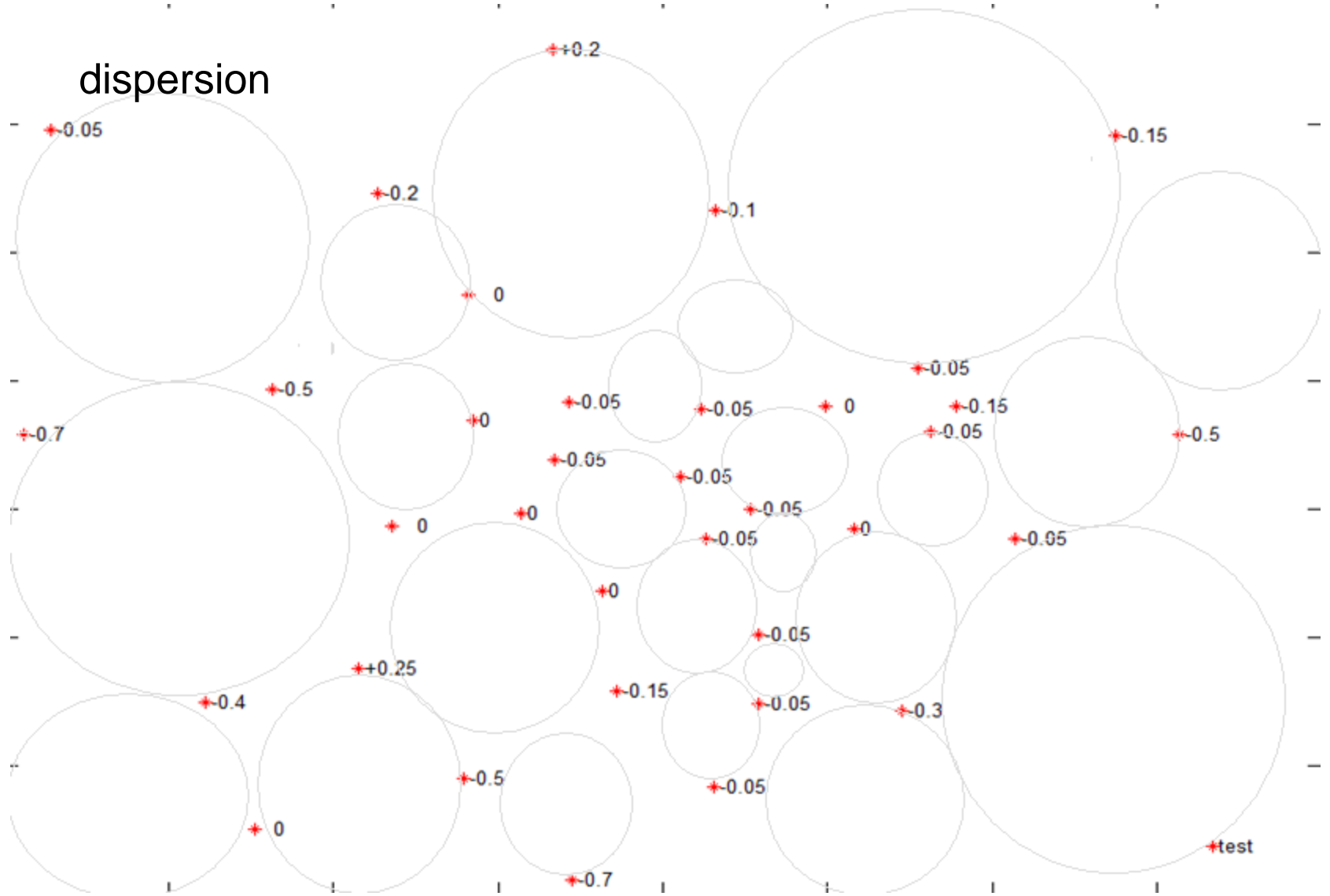
RPM en deux étapes : cartographie sur une surface torique puis dépliement – optimisation par méthode du gradient – l'algorithme cherche à minimiser une énergie globale (méthode de Newton-Raphson) – déformation possible à l'intérieur de la surface de projection

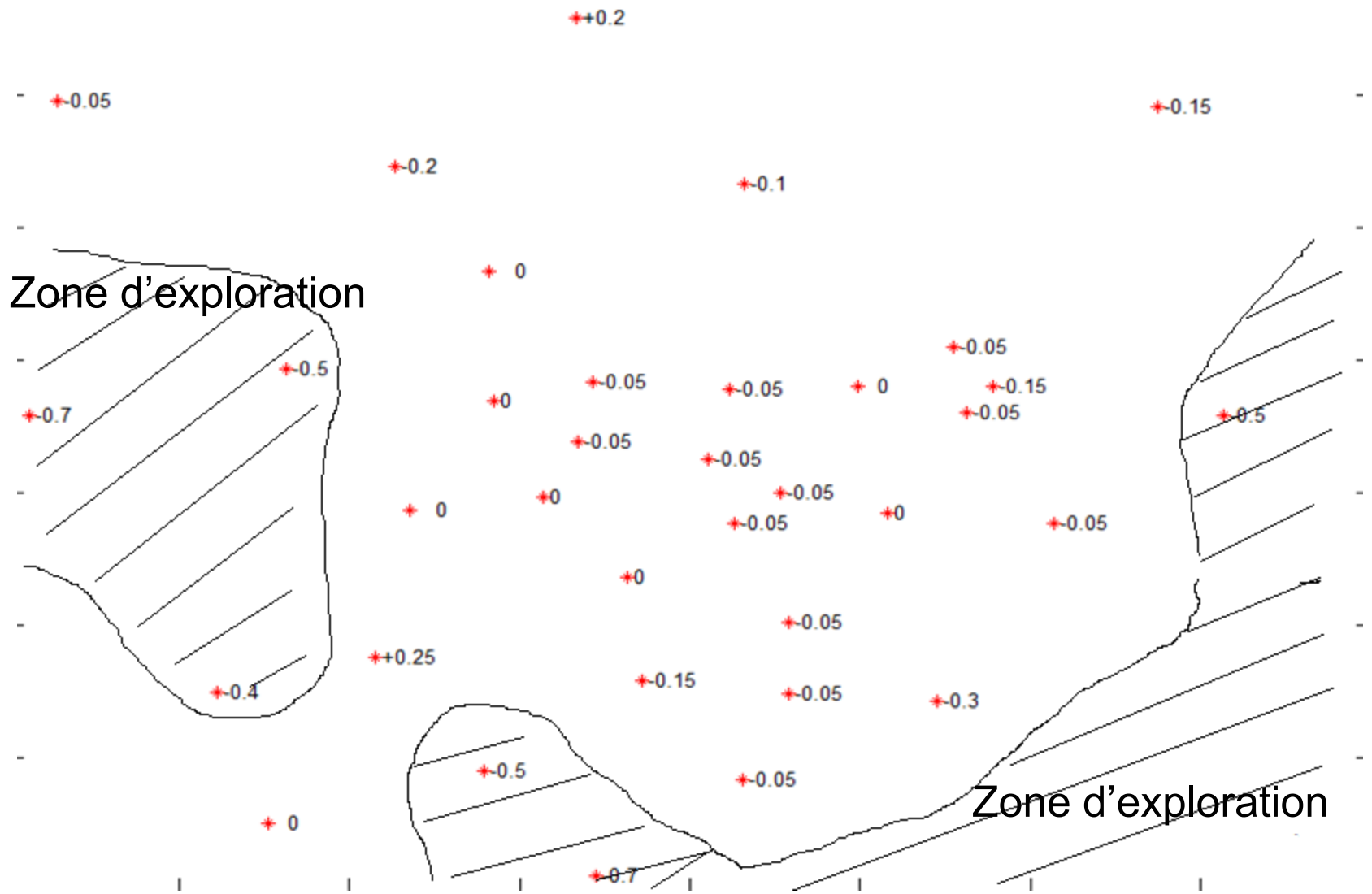
ACC variante non-linéaire de MDS- efficace dans le cadre du dépliement des données d'une structure complexe – fonction Stress augmentée d'un terme F –

Kohonen préserve la topologie, quantification vectorielle sous contrainte de voisinage dimension fixe et rigide – algorithme sensible à la densité de probabilité

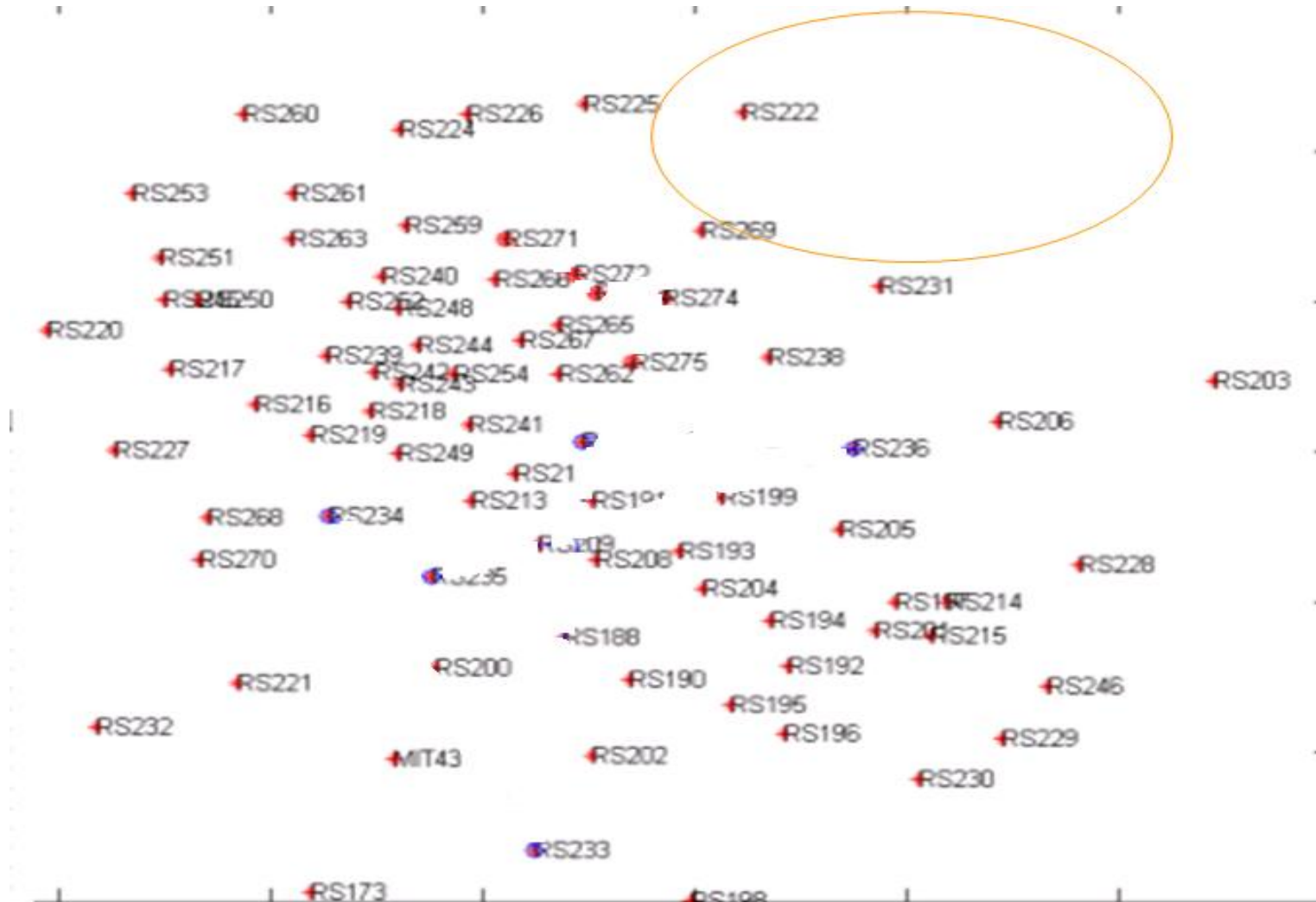


dispersion



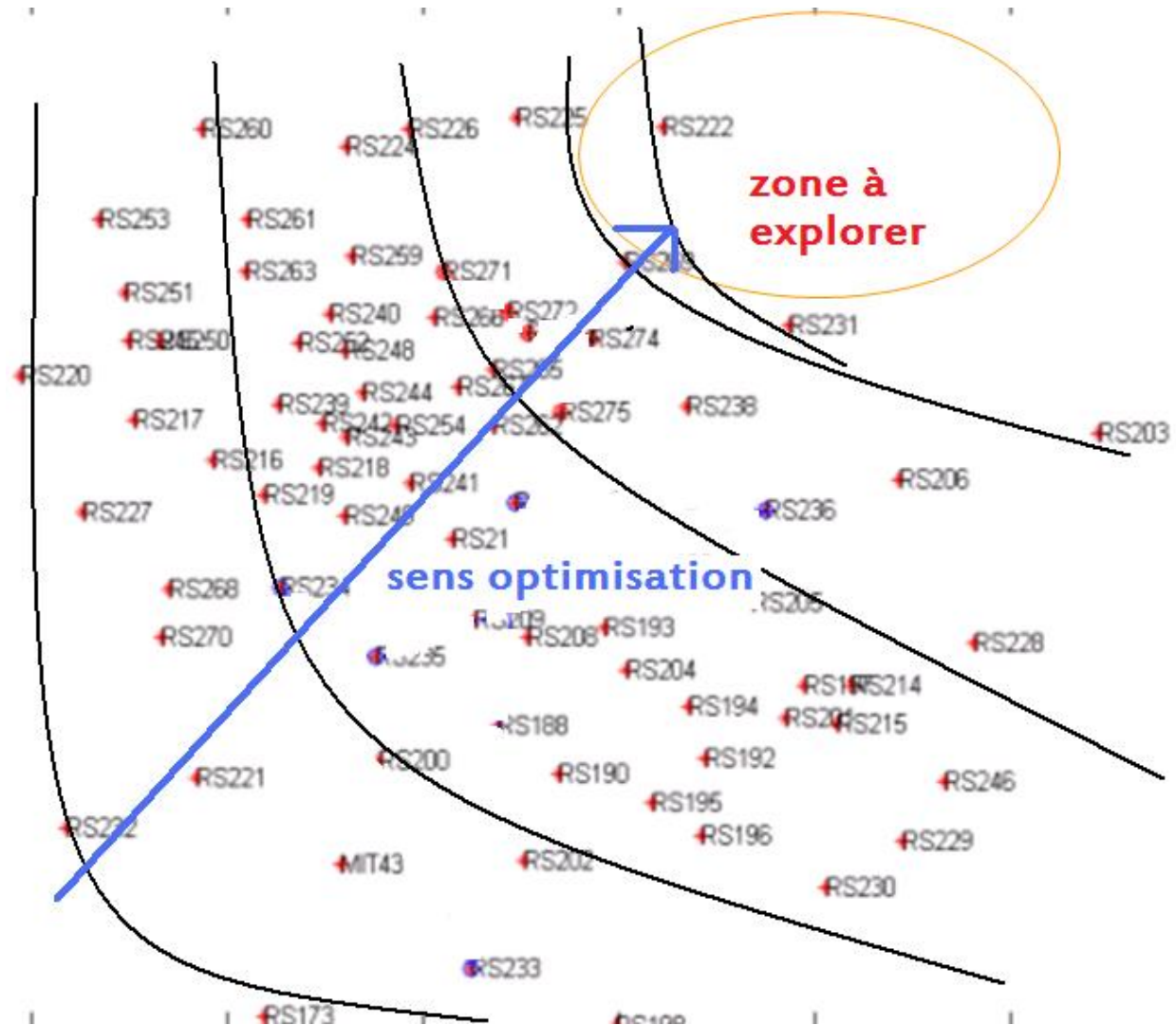


Autre cas de mélanges : après projection ACC



Mélanges à 22 composants sous contraintes doubles

Autre cas de mélanges : après projection ACC



Mélanges à 22 composants sous contraintes doubles



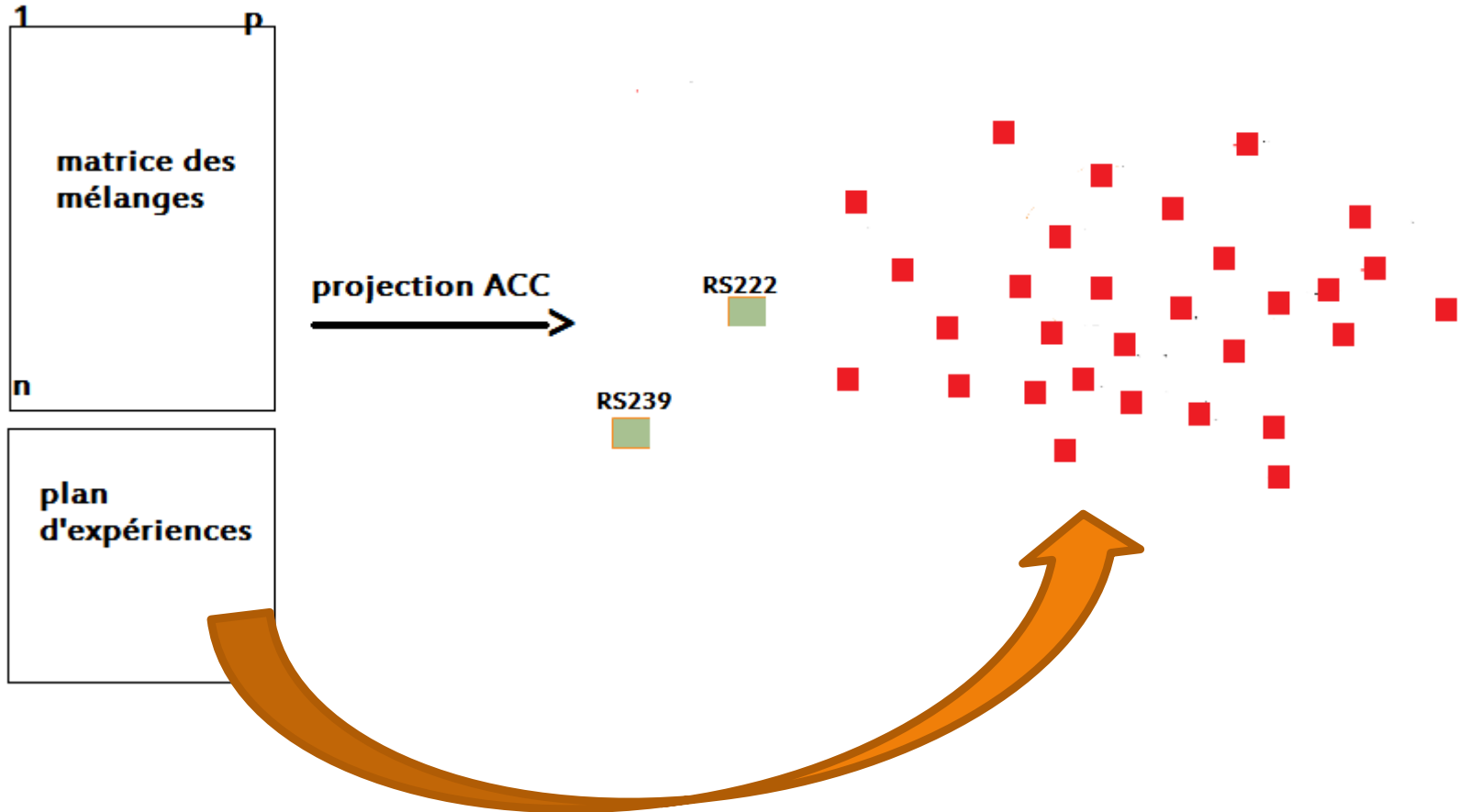
Question 1: comment explorer cette zone

- Structure
- optima locaux

Question 2:

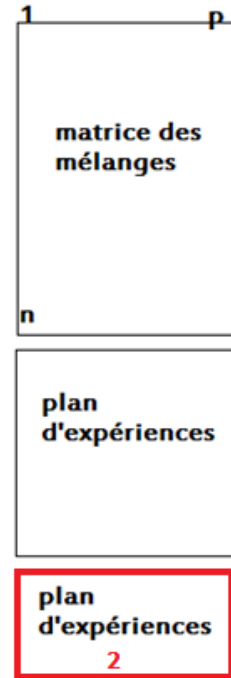
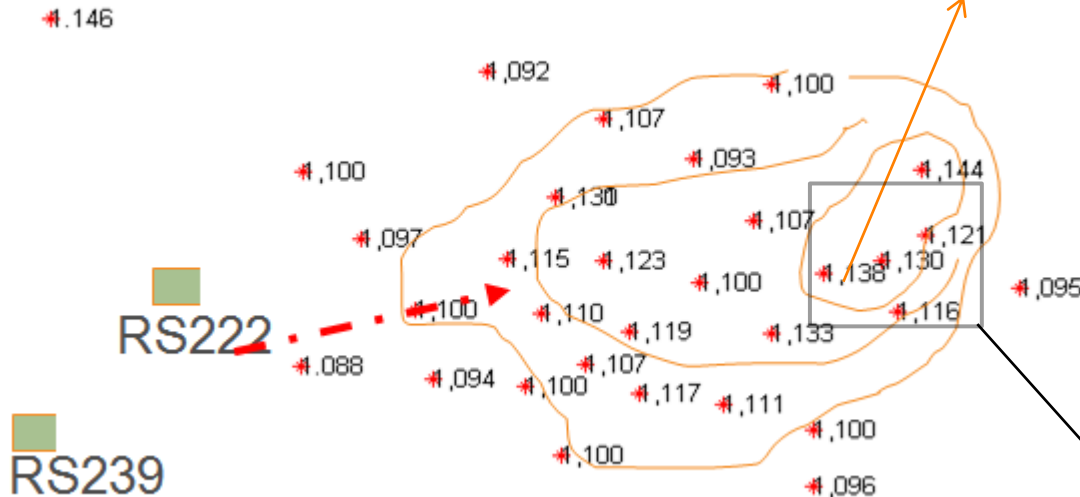
On peut établir une distribution spatiale de n points en 2D mais on ne sait pas revenir aux dimensions n Dim

Proposition C-A S



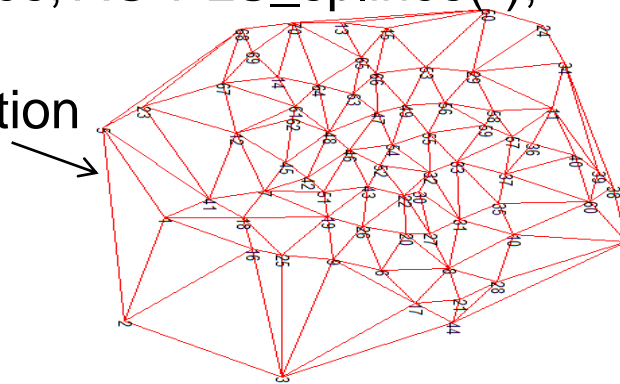
Modélisation

On connaît les coordonnées de ces points



Régression polynomiale, SVM, NN,
PLS, B-splines, AG-PLS_splines(*),
Krigage,
par triangulation

....



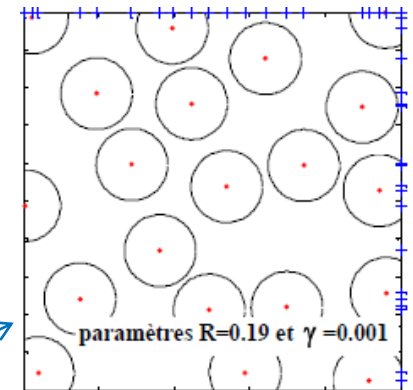
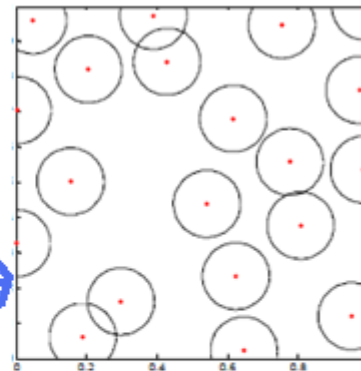
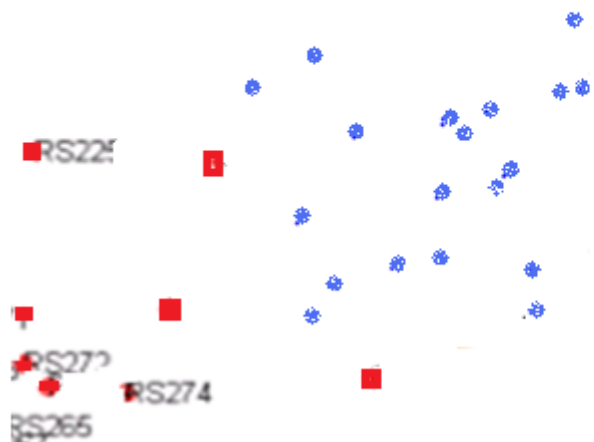
Triangles de Delaunay

Triangles de Delaunay

A(*) G-PLS_splines : les AG permettent d'optimiser
Automatiquement les paramètres des transformations
B-splines.

Si on pouvait faire du dépliement inverse...

- Plan de Strauss



- un rayon trop petit engendre une distribution sans interaction mais avec de nombreuses lacunes (distribution uniforme),
- un rayon trop grand conduit à une distribution avec des agglomérats (clusters).

il est plus aisé de générer une distribution répondant au critère de remplissage de l'espace avec un paramètre de répulsion faible. Cependant, un paramètre de répulsion trop faible contraint généralement les points à être alignés.

il est délicat de bien fixer le paramètre de répulsion afin de trouver un compromis entre un remplissage de l'espace optimal et une répartition de points uniforme

Le rayon d'interaction est le paramètre le plus sensible à régler.

Extensions : processus ponctuel de Gibbs, plans de Strauss-Gibbs avec ou sans contraintes

Répartition spatiale : géométrie algorithmique

La géométrie algorithmique recouvre un certain nombre de méthodes qui sont basées sur la partition de l'espace à partir de la position des points étudiés : ceci revient à caractériser pour chaque point sa zone d'influence et les relations avec les voisins.

Parmi les techniques les plus connues : diagramme de Voronoi et graphes de Delaunay.

Il s'agit de construire une triangulation qui relie les points voisins.

On peut également évoquer les graphes de Gabriel, les graphes de voisins relatifs, les arbres de Longueur minimale.

Un arbre est dit maximal s'il contient tous les points.

Un Arbre de Longueur Minimale (ALM), ou Minimal Spanning Tree (MST), est un arbre maximal pondéré dont la longueur totale (somme de toutes ses arêtes) est la plus petite de tous les arbres maximaux. On peut en effet pondérer un arbre en affectant un poids à chaque arête. Néanmoins, pour un ensemble de points donnés, s'il existe plusieurs arbres maximaux, il n'existe qu'un ALM, par définition.

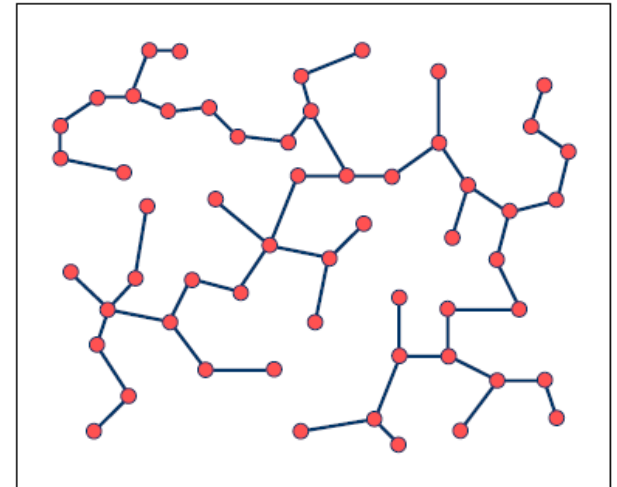
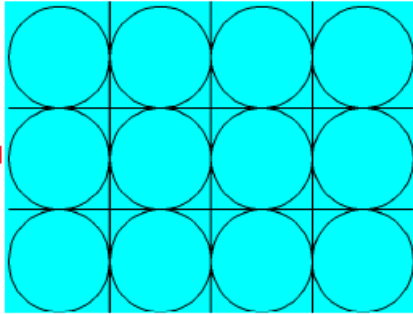
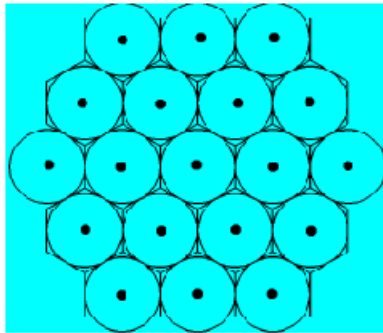


Figure 84. Illustration d'un arbre de longueur minimale



densité

$$\delta = \frac{\text{vol}(\text{disque})}{\text{vol}(\text{carr})} = \frac{\pi}{4}$$

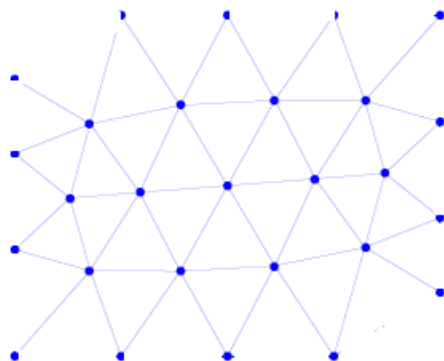


Chaque cellule de Voronoï contient un disque.

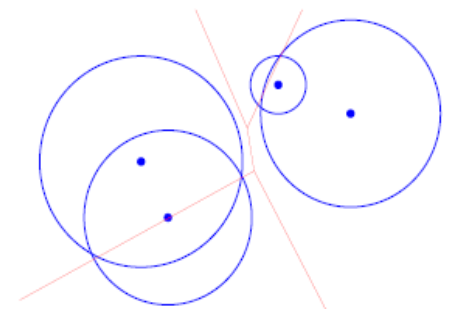
La densité dans la cellule de Voronoï $V(O)$ de centre O est égale à $\frac{\pi}{\text{vol}(V(O))}$.

Extension : diagramme de Laguerre

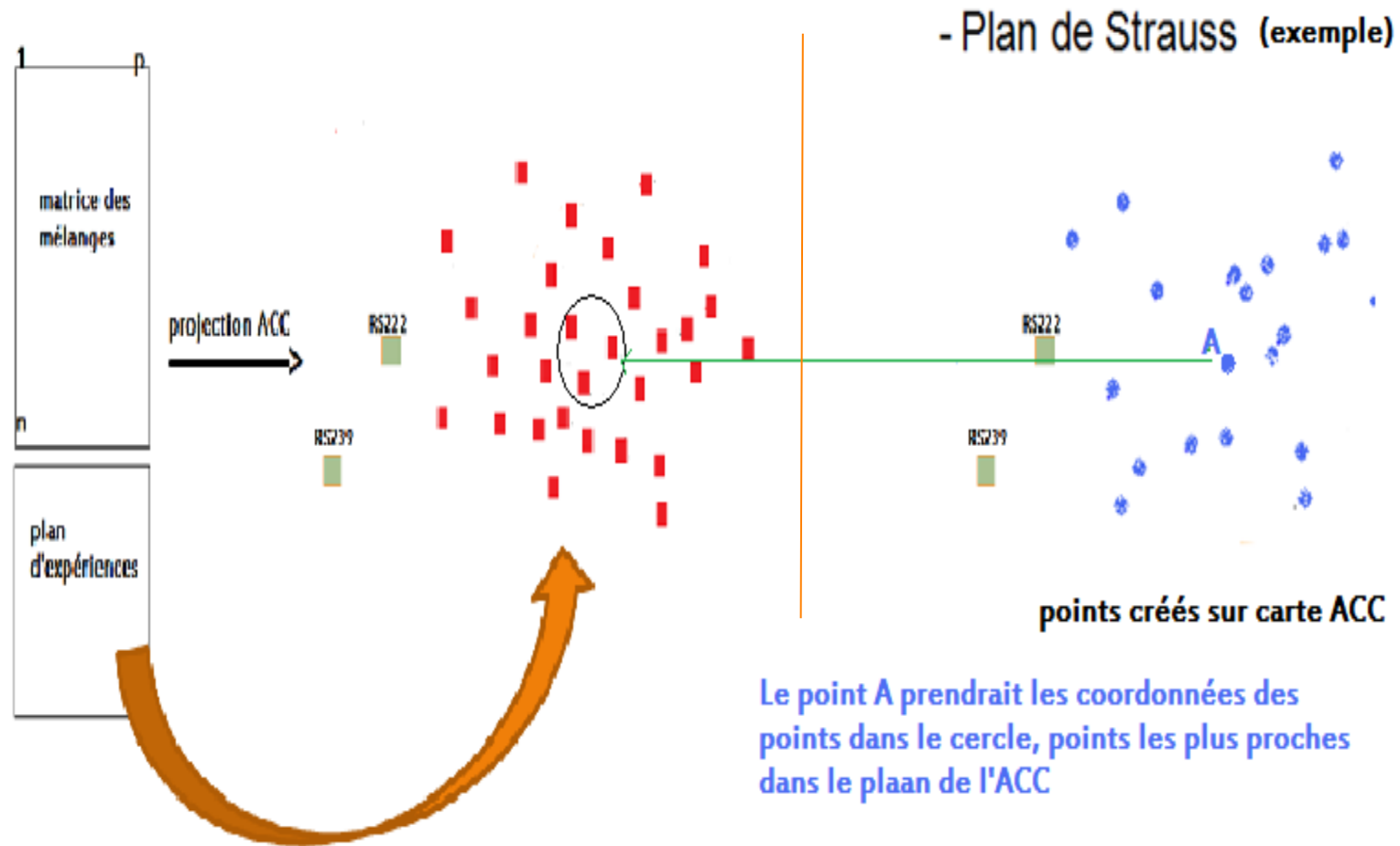
On peut chercher à donner une influence différente À chaque point



Delaunay



Proposition : si l'on ne sait pas faire du dépliement inverse, on procède par comparaison et étude de voisinage :



Autres méthodes possibles de création de points :

- Méthodes sans maillage
 - particules
 - Processus ponctuel
 - Recuit simulé
 - Algorithme génétique
-

Thèse : 'optimisation topologique de structures par algorithmes génétiques'

Par Belkacem SID – Université technologique de Belfort-Montbéliard – 2006

Deux méthodes ont été développées :

- La première propose une approche par A.G basée sur une représentation graphique par courbes de Bézier

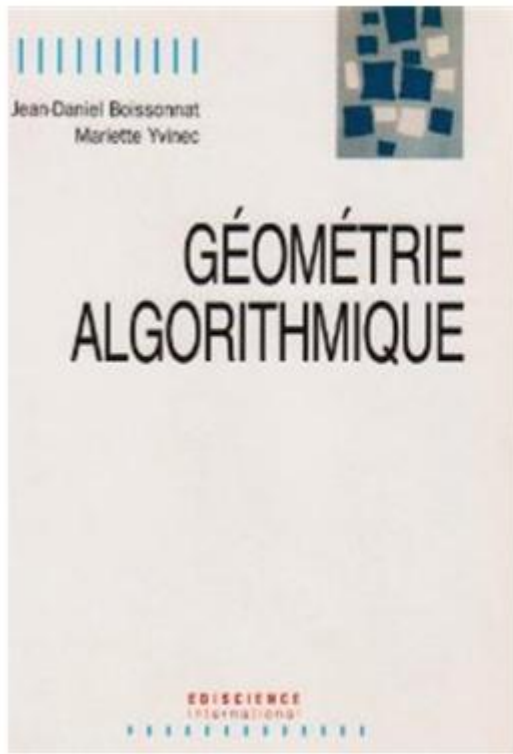
Le principe consiste à construire la structure en utilisant d'abord un squelette que l'on complète.

Le squelette est composé par l'ensemble des courbes cubiques de Bézier reliant les surfaces d'application.

- la deuxième méthode propose une approche par A.G mais basée sur une représentation par adjacence.

Le principe de base de cette représentation est fondé sur le concept de connectivité des éléments finis considérés comme cellules

-



Géométrie algorithmique

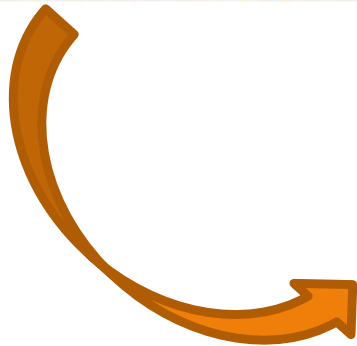
[Jean-Daniel Boissonnat](#) (Auteur),

[Mariette Yvinec](#) (Auteur)

Un petit peu
de géométrie algorithmique

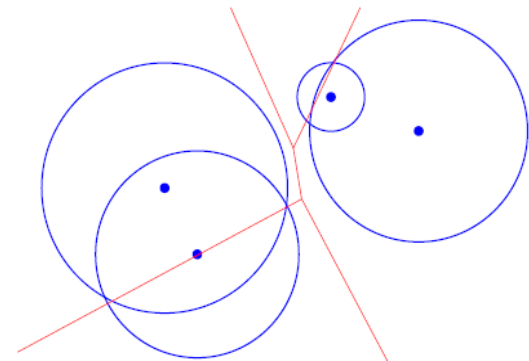
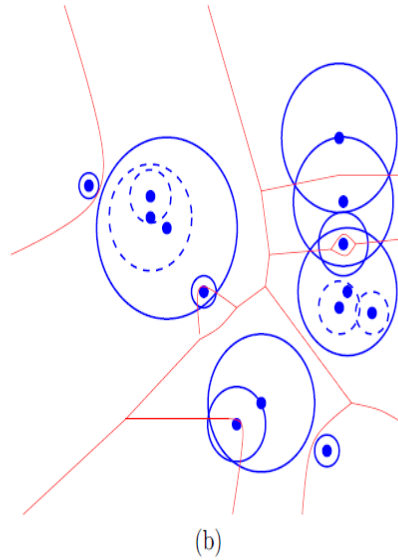
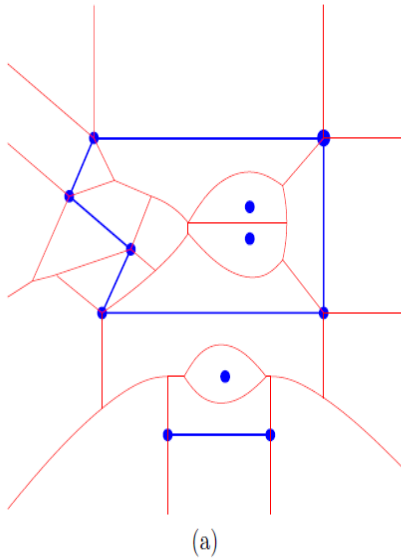
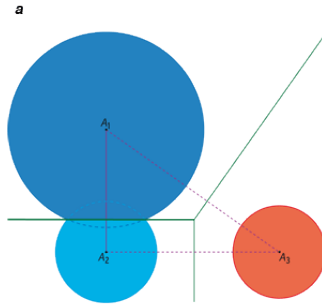
Franck Hétroy

http://evasion.imag.fr/Membres/Franck.Hetroy/Teaching/GeoAlgo/poly_geoalgo.pdf



CHAPITRE 18 : Métriques non euclidiennes

Diagramme de puissance

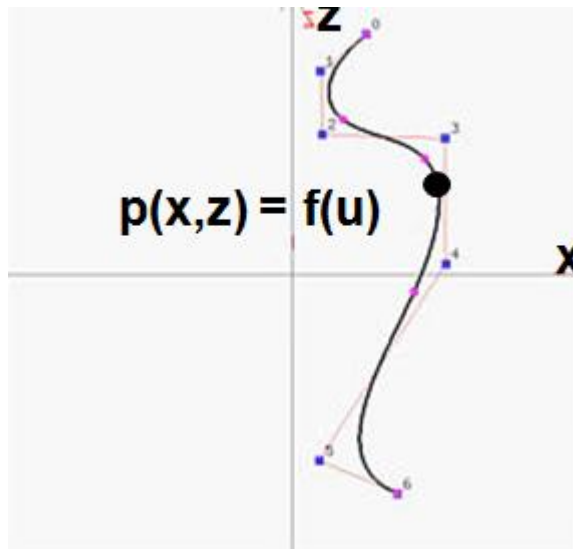


(b) Diagramme de Laguerre.

FIG. 5.1 – (a) Diagramme de Voronoï de points et segments. (b) Diagramme de Voronoï à poids additifs. Les cercles représentent les poids respectifs associés à chaque germe. Certains germes (ceux dont le cercle est en pointillés) ont une cellule associée vide.

Modélisation géométrique Primitives simples

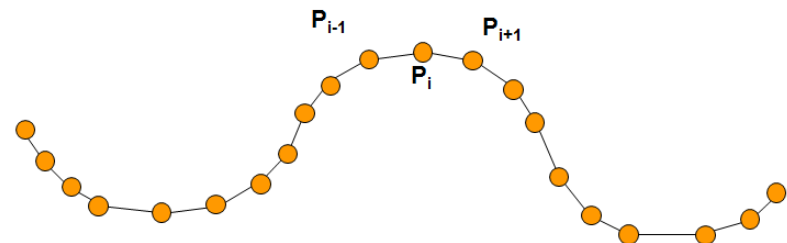
On débute à partir d'une représentation plus ou moins grossière



Repère de Frénet

Le repère local est calculé à partir de l'équation du profil si elle existe, ou à partir de l'approximation discrète de la trajectoire

Evaluation de la tangente à partir de la représentation discrète de la trajectoire (liste de points ou poly-line):



1

N° d'ordre : 2004- 39

ECOLE DOCTORALE 432

École Nationale Supérieure d'Arts et Métiers
Centre de Paris

THÈSE

présentée pour obtenir le grade de

DOCTEUR
de
**L'ÉCOLE NATIONALE SUPÉRIEURE
D'ARTS ET MÉTIERS**

Spécialité : Mécanique

par

Julien YVONNET

**NOUVELLES APPROCHES SANS MAILLAGE BASEES SUR LA METHODE DES
ELEMENTS NATURELS POUR LA SIMULATION NUMERIQUE DES PROCÉDES
DE MISE EN FORME**

Soutenance prévue le 14 décembre 2004 devant le jury composé de

MM.	J.-L. CHENOT	Professeur, ENSMP, Sophia Antipolis	Examinateur
	A. POITOU	Professeur, l'ECN, Nantes	Rapporteur
	P. VILLON	Professeur, UTC, Compiègne	Rapporteur
	P. DEVALAN	Directeur de l'agence de programme, CETIM	Examinateur
	F. CHINESTA	Professeur, l'ENSAM, Paris	Directeur de thèse
	P. LORONG	Maître de Conférences, l'ENSAM, Paris	Codirecteur de thèse
	E. CUETO	Maître de Conférences, Université de Saragosse	Invité

2



Thèse

*présentée pour obtenir le titre de
Docteur de l'Institut National Polytechnique de Toulouse
Spécialité : Mécanique de Fluides*

Développement d'une méthode de simulation d'écoulements à bulles et à gouttes.

par

Thomas BONOMETTI

Thèse soutenue le 29 novembre 2005 devant le jury composé de :

A. BERLEMONT	Membre
A. CARTELLIER	Rapporteur
J.F. DOMGIN	Membre invité
P. GARDIN	Membre
C. GOURDON	Président
P. LE QUÉRÉ	Rapporteur
J. MAGNAUDET	Directeur des travaux de recherche

N° d'ordre :

Exemple :

- ✓ Programme Matlab
- ✓ Déploiement inverse

18

Formulation essences



Matrice des distances

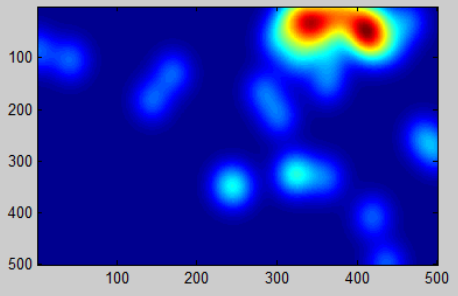
ACC



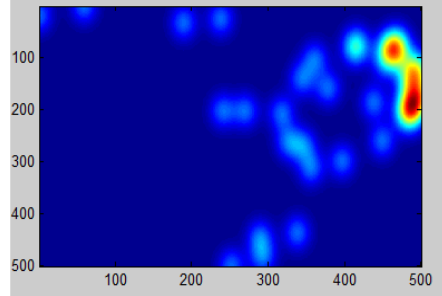
Plan de projection sans dimension

37

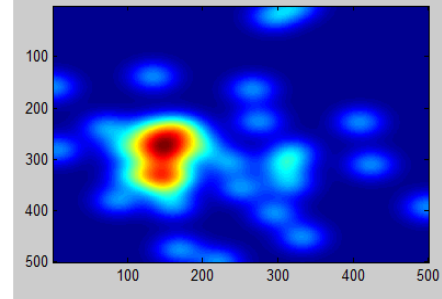
APP : On utilise les 2 premiers axes de l'ACP- input



APP : a partir des distances input

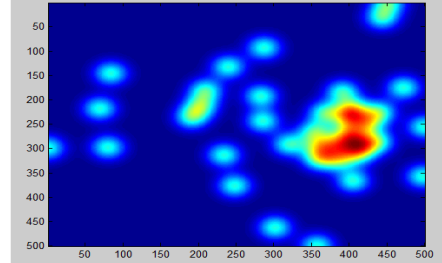


APP : On utilise les coordonnées de l'ACC- output



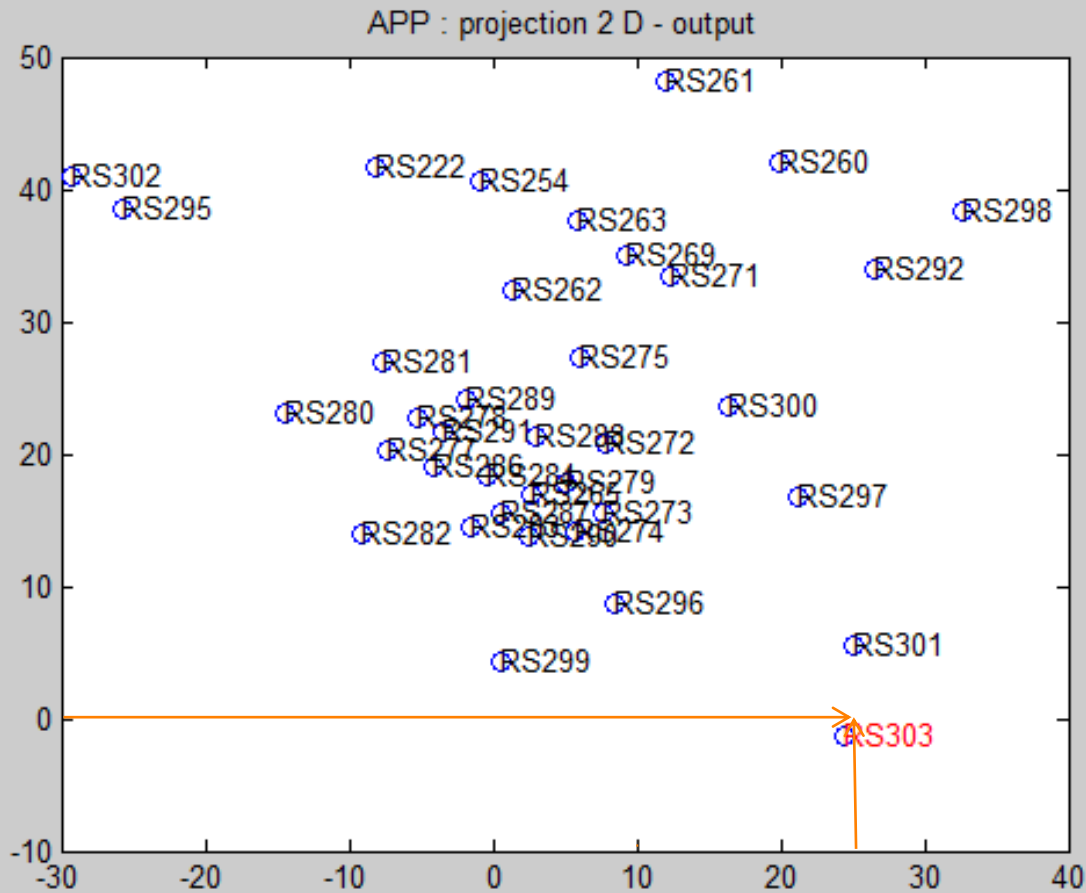
Fonction Potentiel
Pour visualiser les structures
Des matrices

Potentiel sur les distances predites des INPUT



Après repliement
(transformation inverse)





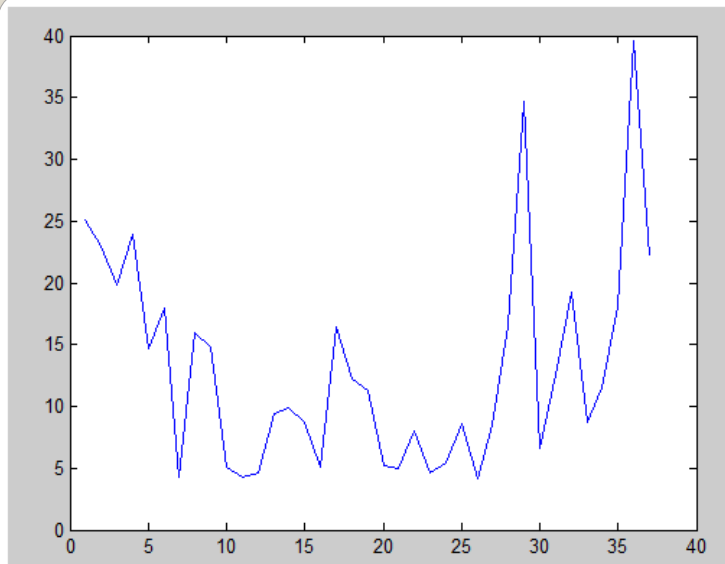
Positionnement des essences après ACC (analyse en composantes curvilignes)

On cherche à connaître la Formule de ce point

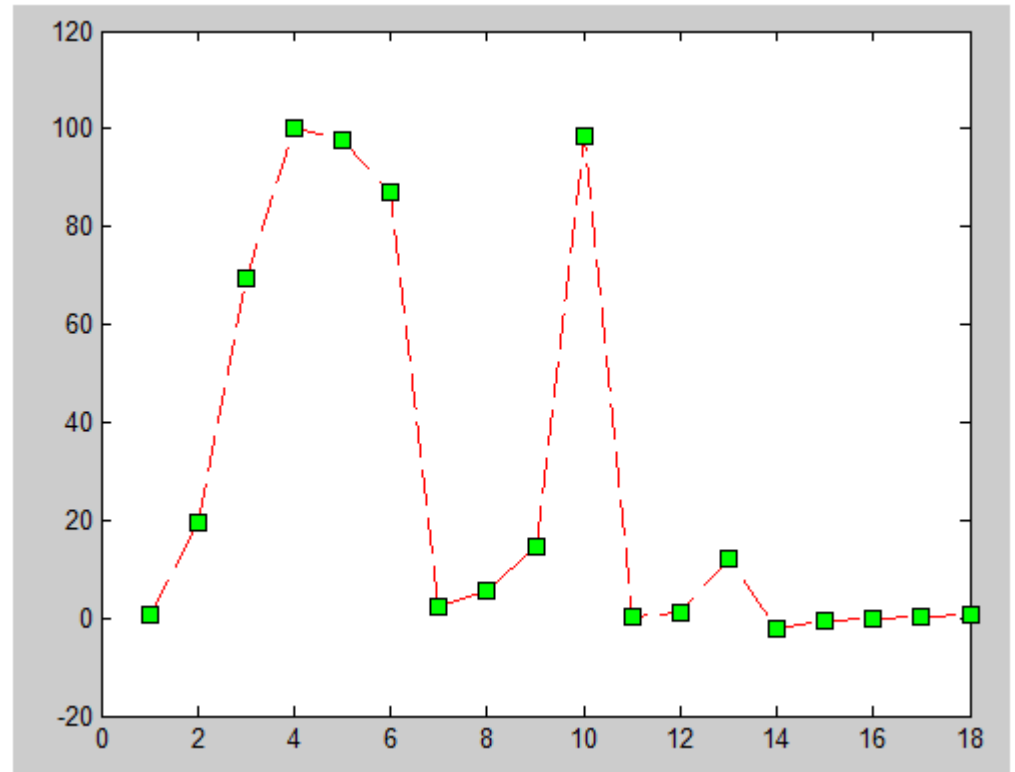
Saisir :

Coordonnées du point sur l axe des x : **25**

Coordonnées du point sur l axe des y : **0**



Ce graphique indique les voisins les plus proches



Valeurs calculées des 18 variables du produit inconnu

Calcul de la réponse Y du point test connaissant maintenant ses propriétés

réponse Y

```
-0.15  
-0.05  
-0.5  
-0.7  
-0.05  
-0.05  
-0.05  
-0.3  
-0.15  
-0.05  
-0.05  
0  
0  
-0.05  
-0.05  
-0.05  
-0.05  
-0.15  
-0.5  
-0.05  
0  
-0.05  
-0.05  
0  
0  
0  
0  
-0.7  
0.25  
-0.2  
-0.4  
-0.5  
-0.1  
0  
0.2  
0.35  
0
```

Valeur réelle du point test

2 modèles :

```
=====
% Calcul du vecteur reponse Y du nouveau point : par PLS
=====
delete X;delete yn;delete yfit;delete Xt; delete y_test;
X=M_app;
yn=cas_essence_y;
[XL,y1,XS,YS,beta,PCTVAR,MSE,stats] = plsregress(X,yn,q);
yfit = [ones(size(X,1),1) X]*beta;

%=== test =====
Xt=M_test3;
y_test = [ones(size(Xt,1),1) Xt]*beta;
if y_test <0
    y_test=0
end
y_test

% Calcul du vecteur reponse Y du nouveau point : par Triangulation
=====
Xcal=M_app;
Ycal=cas_essence_y;
Xtest=M_test3;
Ytest=1;
prep=0,opt=2,pc=2,dim=2
[REPONSE, REPONout,TRI,t2,outliers,ttout]=dtnew(Xcal,Ycal,Xtest,...
    Ytest,prep,opt,pc,dim)
REPONSE
REPONout
valeur=REPONSE(:,2:2)
SYtest=sort(Ytest)
REP=[REPONSE;REPONout]
```

Calcul de la réponse Y du point test connaissant maintenant ses propriétés

réponse Y

2 modèles :

```
-0.15  
-0.05  
-0.5  
-0.7  
-0.05  
-0.05  
-0.05  
-0.3  
-0.15  
-0.05  
-0.05  
0  
0  
-0.05  
-0.05  
-0.05  
-0.05  
-0.15  
-0.5  
-0.05  
0  
-0.05  
-0.05  
0  
0  
0  
0  
-0.7  
0.25  
-0.2  
-0.4  
-0.5  
-0.1  
0  
0.2  
0.35  
0
```

Calcul du vecteur reponse Y du nouveau point : par PLS

`y_test = 0.`

Calcul du vecteur reponse Y du nouveau point : par Triangulation

`valeur = 0`

Valeur réelle du point test



Questions ?