

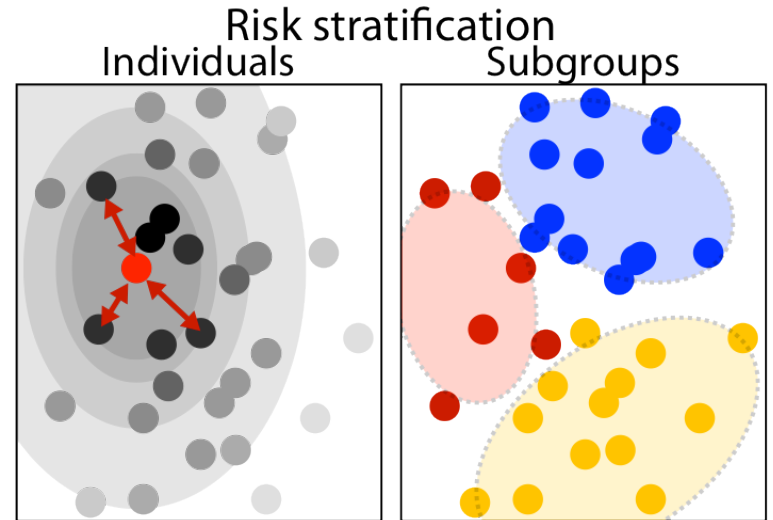
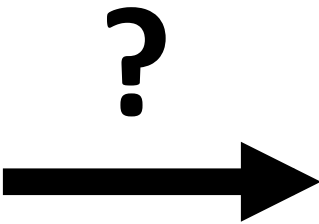
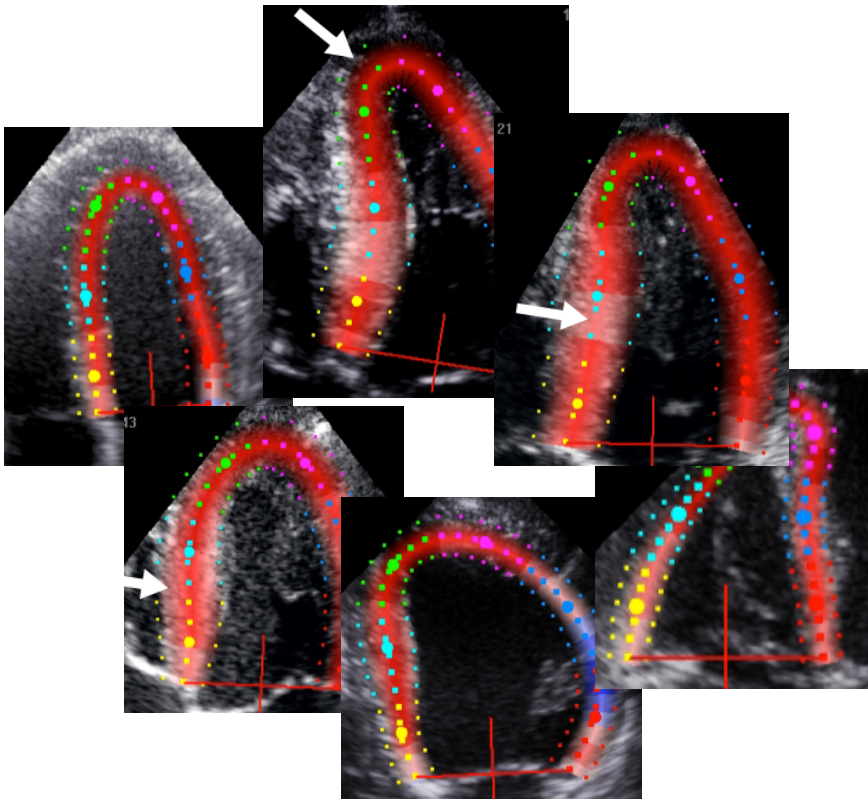
Fusion de données d'imagerie médicale: réduction de dimensionnalité et apprentissage statistique

Nicolas Duchateau

CREATIS, Université Lyon 1, FR

Imagerie médicale (cardiaque): enjeux

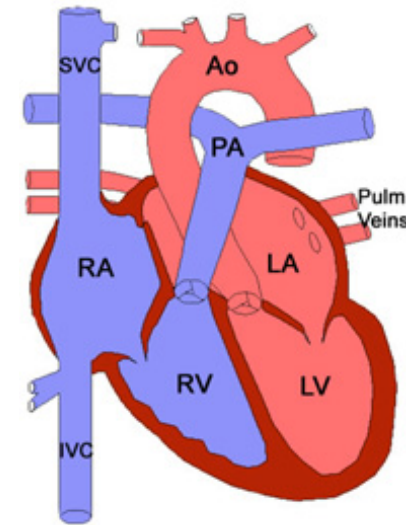
- Comment **représenter** une population ?
- Comment stratifier le **risque** ?



Imagerie médicale (cardiaque): enjeux

- Comment **représenter** une population ?
- Comment stratifier le **risque** ?

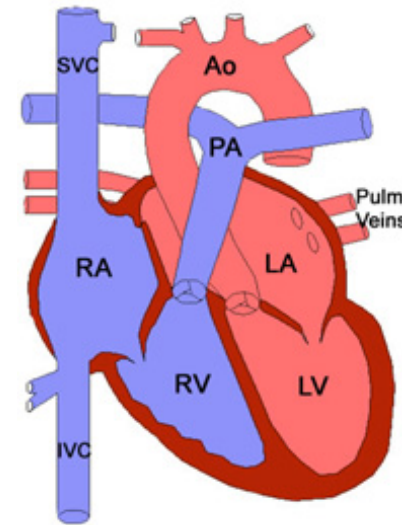
Application	Modalité		Sous-organe
Resynchronisation	2D / 3D	ECHO + CARTO	LV
Infarctus	3D	ECHO + MR	LV
Surcharge en pression (PH)	3D	ECHO	RV
Ejection préservée	2D + TDI	ECHO	LV + LA/RA
Surcharge en volume (Fallot + ASD)	3D	ECHO	RV
Cardiomyopathie hypertrophique (HOCM)	2D	ECHO	LV
Athlètes / effort	2D	ECHO	LV / RV / LA / RA



Imagerie médicale (cardiaque): enjeux

- Comment **représenter** une population ?
- Comment stratifier le **risque** ?

Application	Modalité		Sous-organe
Resynchronisation	2D / 3D	ECHO + CARTO	LV
Infarctus	3D	ECHO + MR	LV
Surcharge en pression (PH)	3D	ECHO	RV
Ejection préservée	2D + TDI	ECHO	LV + LA/RA
Surcharge en volume (Fallot + ASD)	3D	ECHO	RV
Cardiomyopathie hypertrophique (HOCM)	2D	ECHO	LV
Athlètes / effort	2D	ECHO	LV / RV / LA / RA



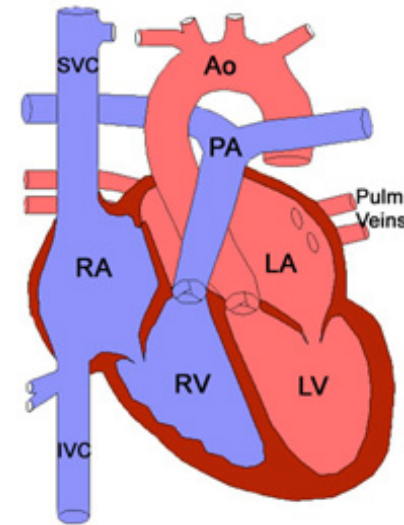
Conditionné par les données initiales:

1 point de vue,

Imagerie médicale (cardiaque): enjeux

- Comment **représenter** une population ?
- Comment stratifier le **risque** ?

Application	Modalité		Sous-organe
Resynchronisation	2D / 3D	ECHO + CARTO	LV
Infarctus	3D	ECHO + MR	LV
Surcharge en pression (PH)	3D	ECHO	RV
Ejection préservée	2D + TDI	ECHO	LV + LA/RA
Surcharge en volume (Fallot + ASD)	3D	ECHO	RV
Cardiomyopathie hypertrophique (HOCM)	2D	ECHO	LV
Athlètes / effort	2D	ECHO	LV / RV / LA / RA



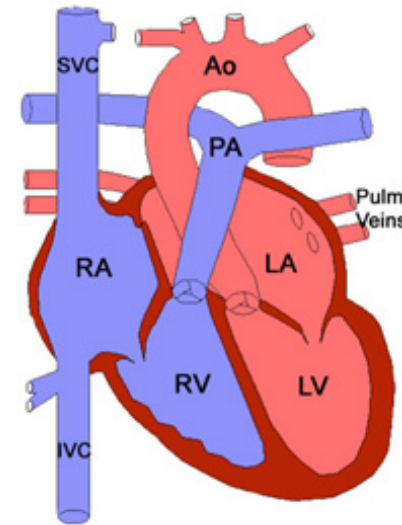
Conditionné par les données initiales:

1 point de vue, **plus ou moins complexe,**

Imagerie médicale (cardiaque): enjeux

- Comment **représenter** une population ?
- Comment stratifier le **risque** ?

Application	Modalité		Sous-organe
Resynchronisation	2D / 3D	ECHO + CARTO	LV
Infarctus	3D	ECHO + MR	LV
Surcharge en pression (PH)	3D	ECHO	RV
Ejection préservée	2D + TDI	ECHO	LV + LA/RA
Surcharge en volume (Fallot + ASD)	3D	ECHO	RV
Cardiomyopathie hypertrophique (HOCM)	2D	ECHO	LV
Athlètes / effort	2D	ECHO	LV / RV / LA / RA



Conditionné par les données initiales:

1 point de vue, plus ou moins complexe, **contraintes physiologiques**

Imagerie médicale (cardiaque): données

Images

- Niveaux de gris, texture

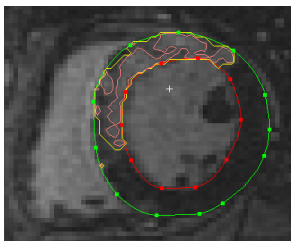
Formes

- **Geometrie (maillage 3D, courbure)**, fibres

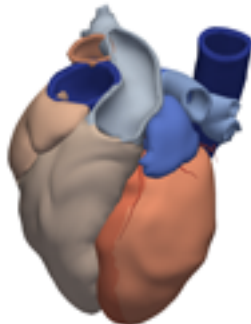
Données fonctionnelles

- Globales: **mesures cliniques**, outcome
- Locales: **mécaniques (mouvement / déformation)**, électriques

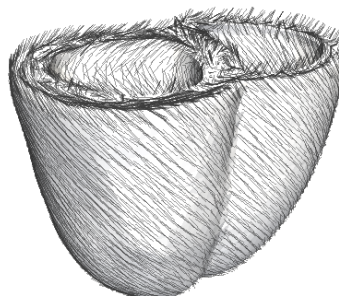
Image brute



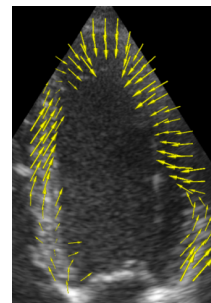
Forme



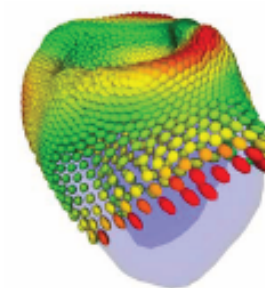
Fibres



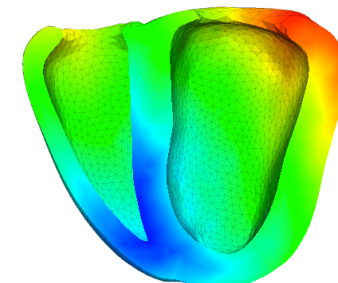
Vitesses



Déformation



Activation
électrique



Imagerie médicale (cardiaque): données

Images

- Niveaux de gris, texture

Formes

- Geometrie (maillage 3D, courbure), fibres

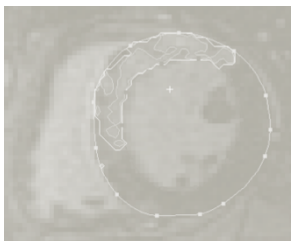
Données fonctionnelles

- Globales: mesures cliniques, outcome
- Locales: mécaniques (mouvement / déformation), électriques

Specificités / contraintes ?

- Physiologie, structure spécifique (variété?)
- 4D (espace + temps) ... ou 5D (longitudinal)
- Haute dimension
- Taille de population: de 100 à 3.000+ sujets

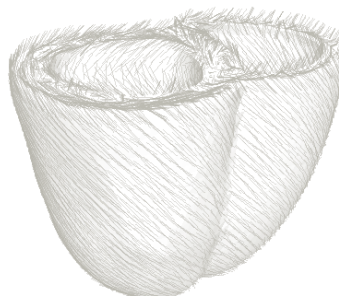
Image brute



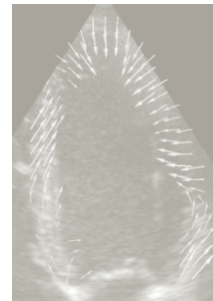
Forme



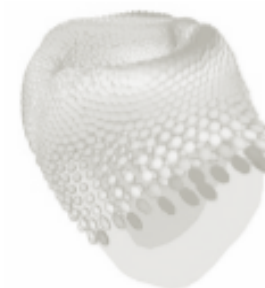
Fibres



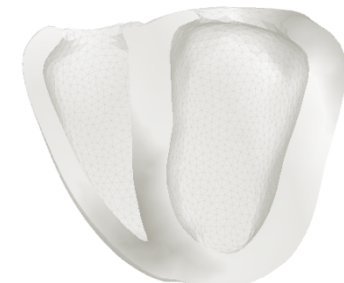
Vitesses



Déformation



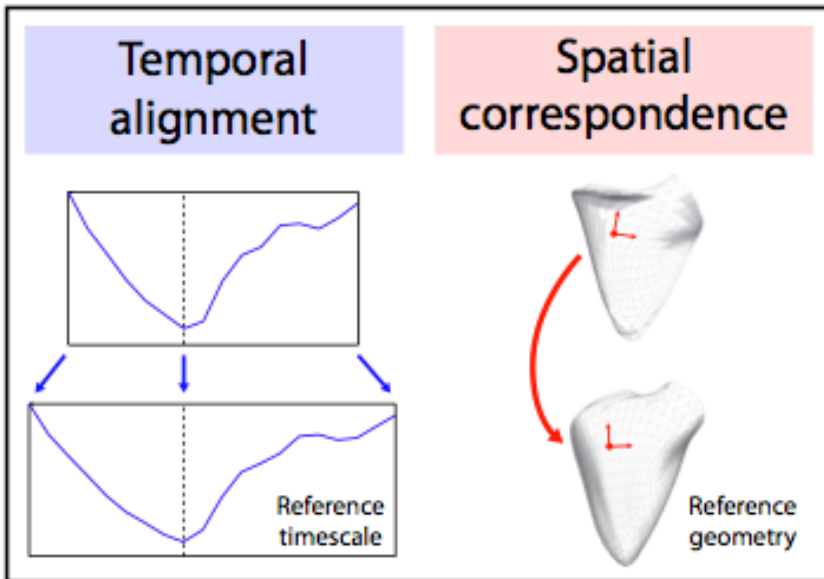
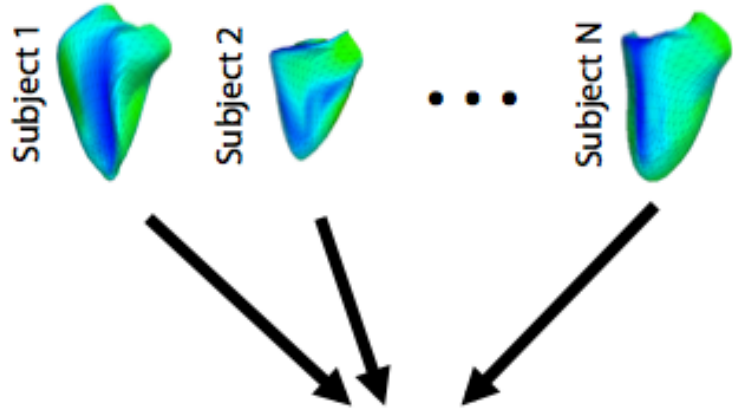
Activation
électrique



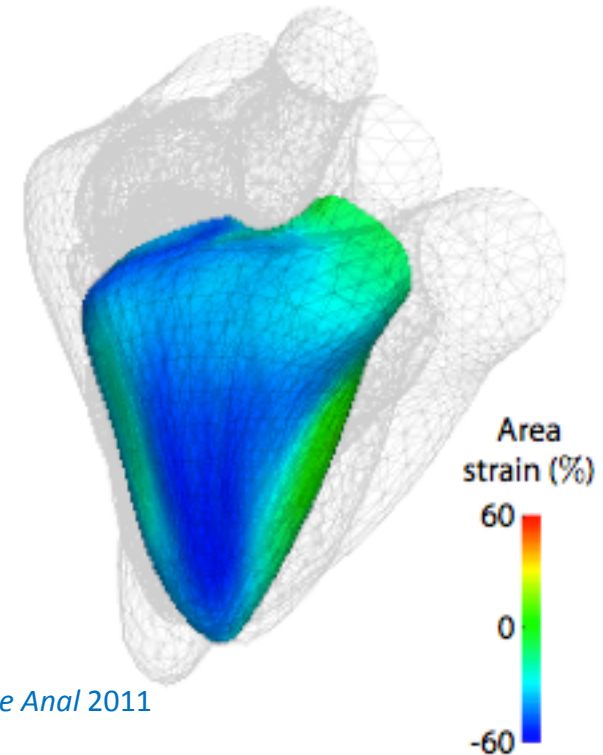
Pré-traitement ?

**Normalisation
intra- et inter-individus**

- échantillonnage
- géométries
- dynamiques



Atlas statistique



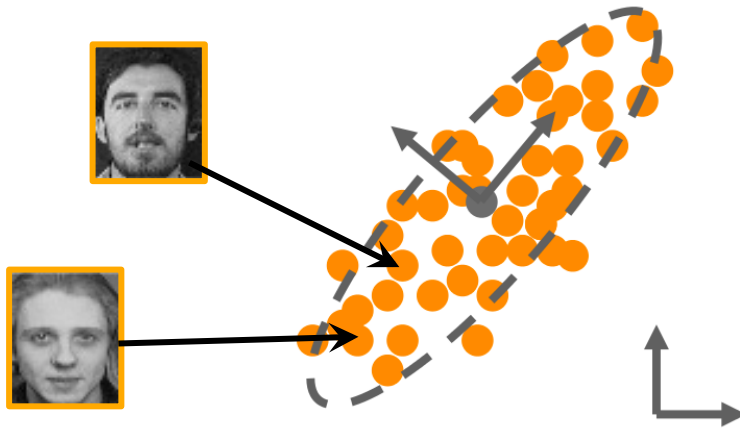
Duchateau et al. *Med Image Anal* 2011
De Craene et al. *ISBI* 2012.

Apprendre à représenter une population

Duchateau et al. *Med Image Anal* 2012
Duchateau et al. *SEE-GSI* 2013

Moyenne & modes de variation

Linéaire = analyse en composantes principales (PCA)



ORL database

diagonalisation de matrice (covariance)
= valeurs propres / vecteurs propres

Apprendre à représenter une population

Duchateau et al. *Med Image Anal* 2012
Duchateau et al. *SEE-GSI* 2013

Moyenne & modes de variation

Linéaire = analyse en composantes principales (PCA)



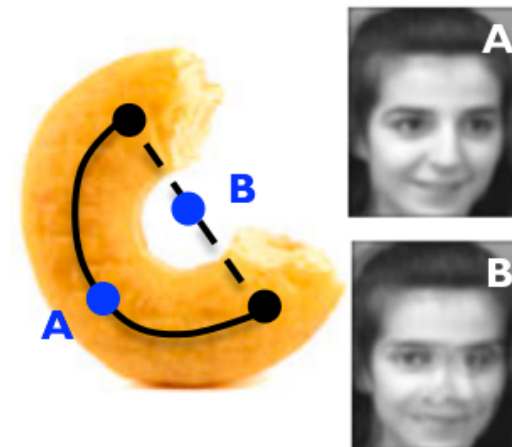
ORL database

diagonalisation de matrice (covariance)
= valeurs propres / vecteurs propres

Non-linéaire !

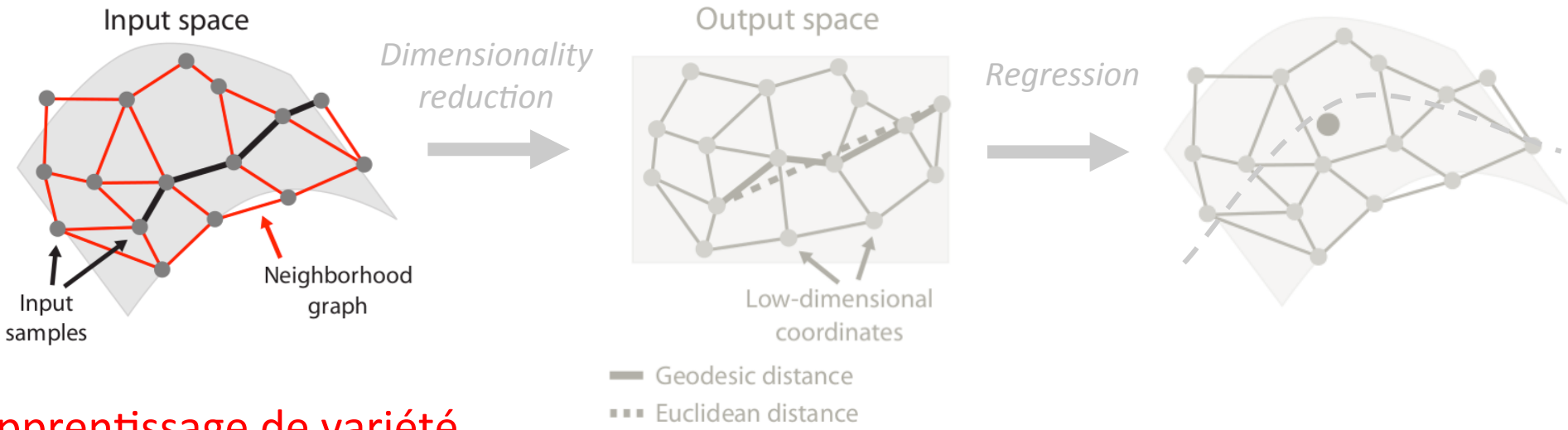


L'apprentissage statistique doit respecter l'espace des données...

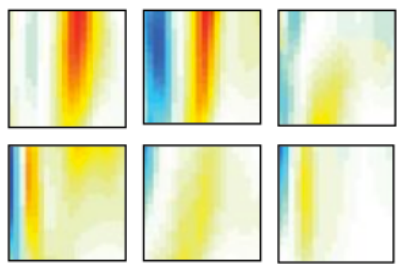


ORL database

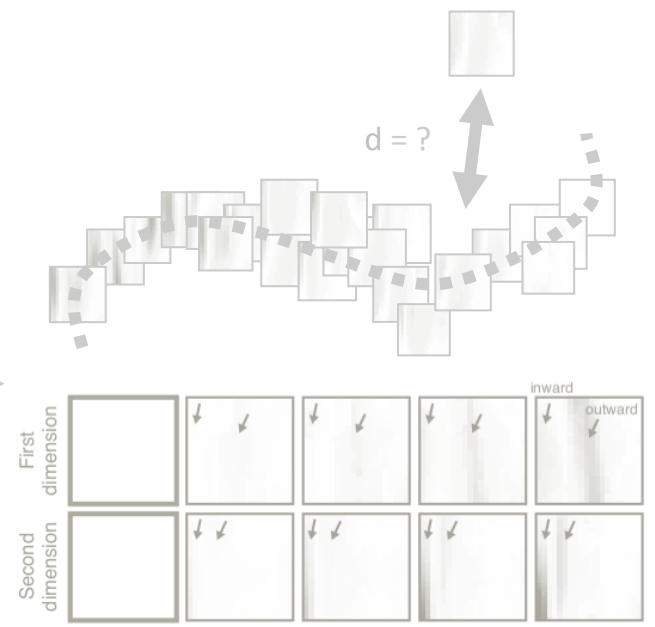
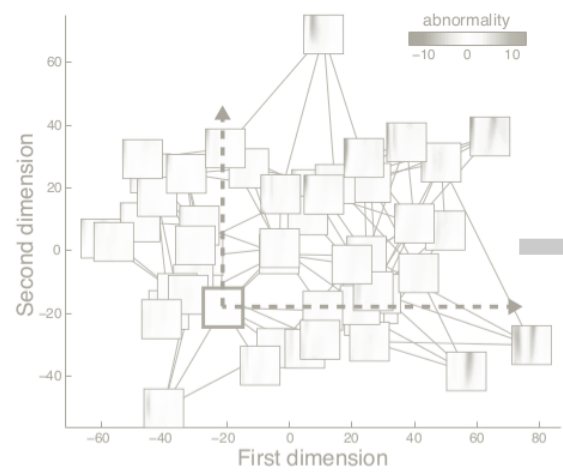
Apprendre à représenter une population Duchateau et al. *Med Image Anal* 2012
Duchateau et al. *SEE-GSI* 2013



= Apprentissage de variété (non-supervisé)

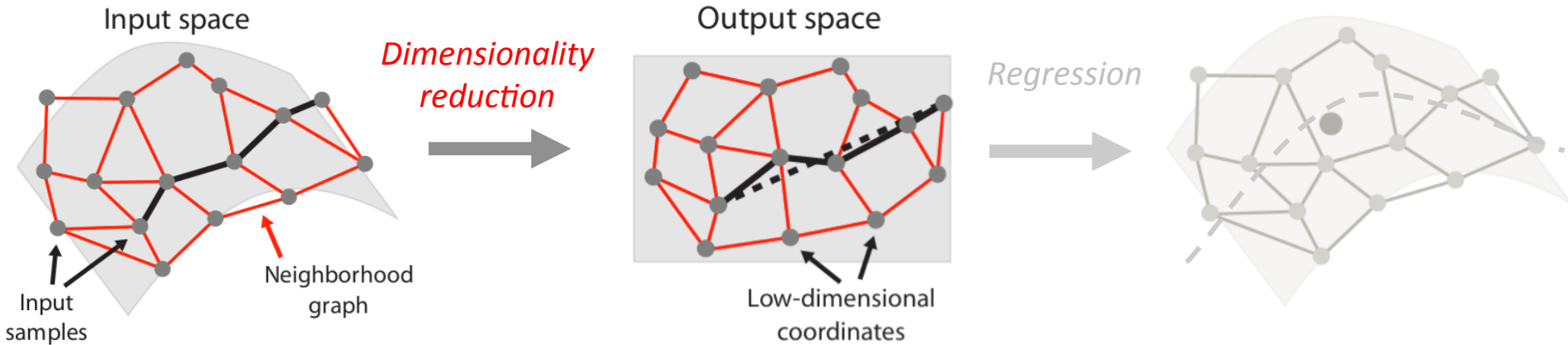


Example: (ab)normal motion patterns

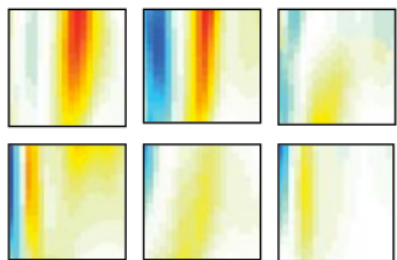


Apprendre à représenter une population

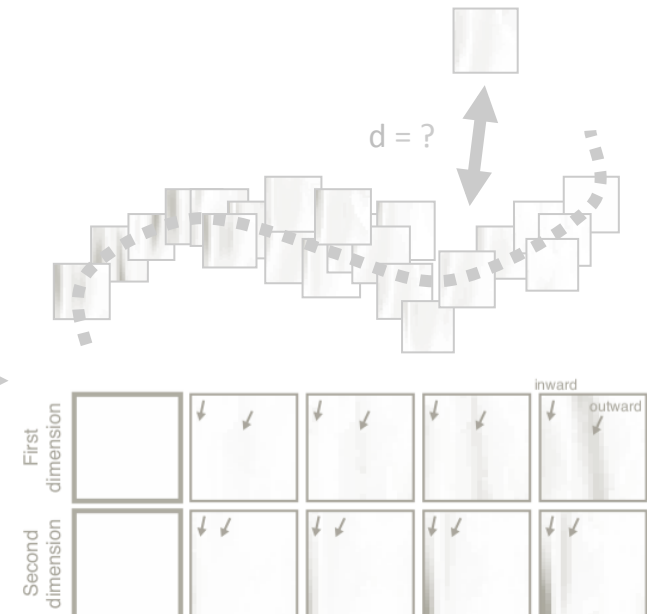
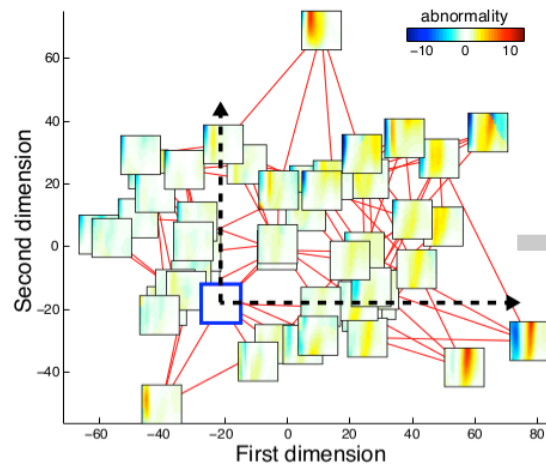
Duchateau et al. *Med Image Anal* 2012
Duchateau et al. *SEE-GSI* 2013



= Apprentissage de variété (non-supervisé)

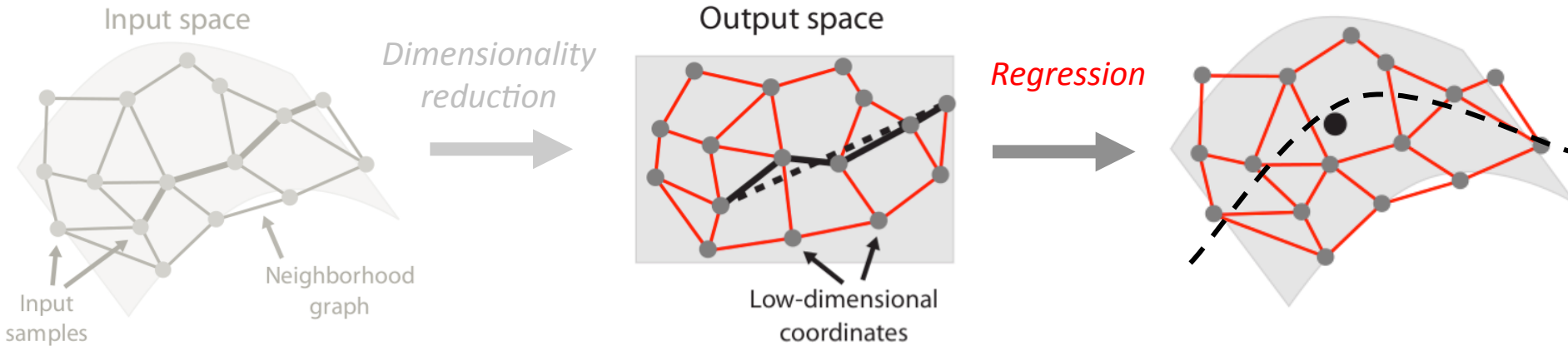


Example: (ab)normal motion patterns



Apprendre à représenter une population

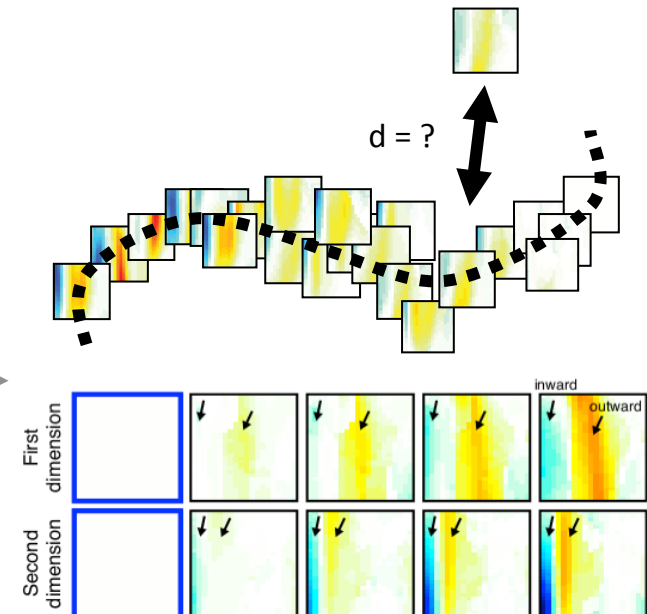
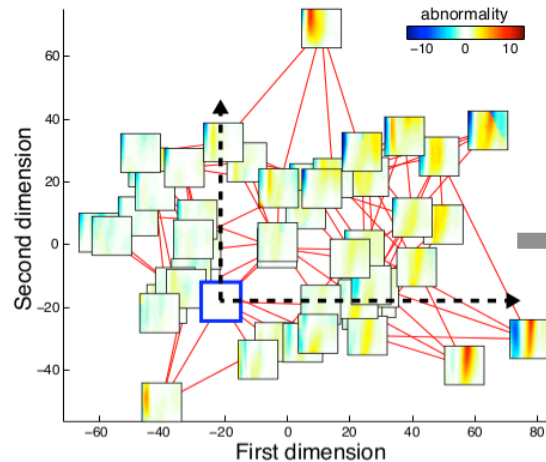
Duchateau et al. *Med Image Anal* 2012
Duchateau et al. *SEE-GSI* 2013



— Geodesic distance
 ... Euclidean distance

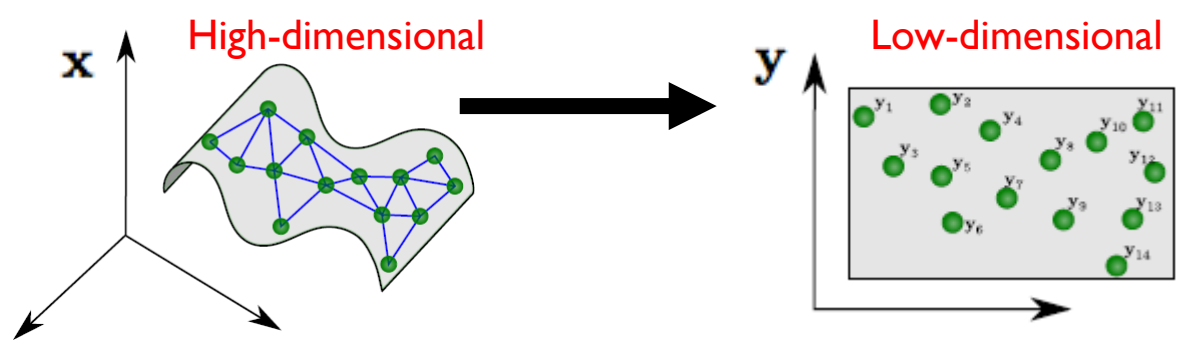
= Apprentissage de variété (non-supervisé)

Example: (ab)normal motion patterns

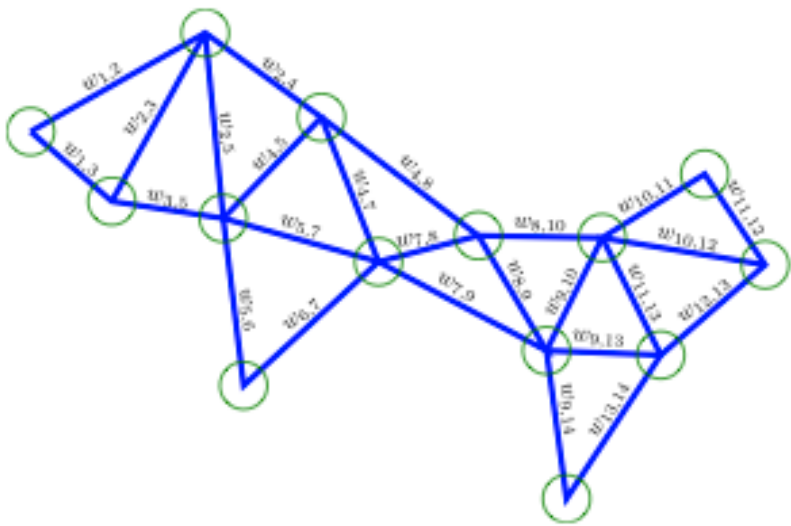


Apprendre à représenter une population: manifold learning / graph embedding

Réduction de dimension: $f : \mathbf{x} \in \mathbb{R}^h \mapsto \mathbf{y} \in \mathbb{R}^l$ $h \gg l$



Atasoy & Mateus, MICCAI 2011



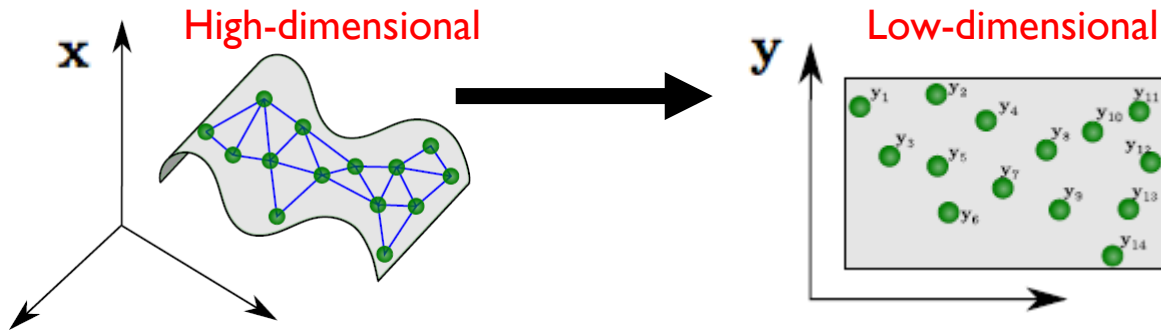
$$W = \begin{bmatrix} 0 & w_{1,2} & w_{1,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ w_{1,2} & 0 & 0 & w_{2,4} & w_{2,5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ w_{1,3} & w_{2,3} & 0 & 0 & w_{3,5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{2,4} & 0 & 0 & w_{4,5} & 0 & w_{4,7} & w_{4,8} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & w_{3,5} & w_{4,5} & 0 & w_{5,6} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{5,6} & w_{6,7} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{4,7} & w_{5,7} & w_{6,7} & 0 & w_{7,8} & w_{7,9} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{4,8} & 0 & 0 & w_{7,8} & 0 & w_{8,9} & w_{8,10} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & w_{7,9} & w_{8,9} & 0 & w_{9,10} & 0 & 0 & 0 & w_{9,13} & w_{9,14} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{8,10} & w_{9,10} & 0 & w_{10,11} & w_{10,12} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{10,11} & 0 & w_{11,12} & w_{11,13} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{10,12} & w_{11,12} & 0 & w_{12,13} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{9,13} & 0 & w_{11,13} & w_{12,13} & 0 & 0 & w_{13,14} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{9,14} & 0 & 0 & 0 & 0 & w_{13,14} & 0 \end{bmatrix}$$

$$\underbrace{\begin{pmatrix} \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \end{pmatrix}}_{T(W)} = \underbrace{\begin{pmatrix} \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \end{pmatrix}}_V \underbrace{\begin{pmatrix} \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \end{pmatrix}}_\Sigma \underbrace{\begin{pmatrix} \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \end{pmatrix}}_{V^T}$$

Apprendre à représenter une population: manifold learning / graph embedding

Réduction de dimension:

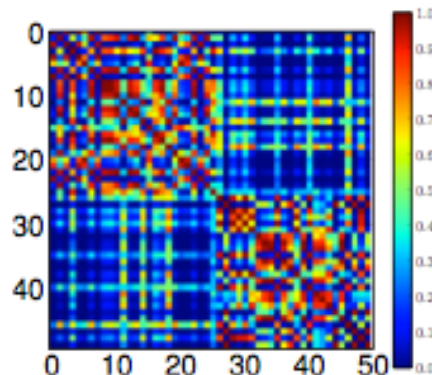
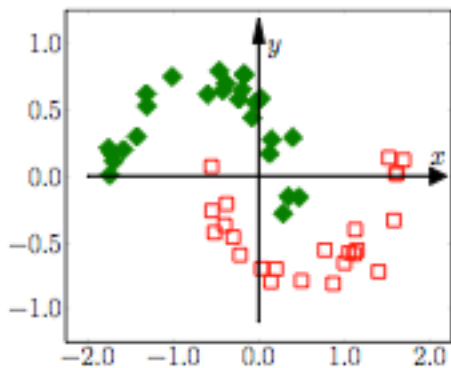
$$f : \mathbf{x} \in \mathbb{R}^h \mapsto \mathbf{y} \in \mathbb{R}^l \quad h \gg l$$



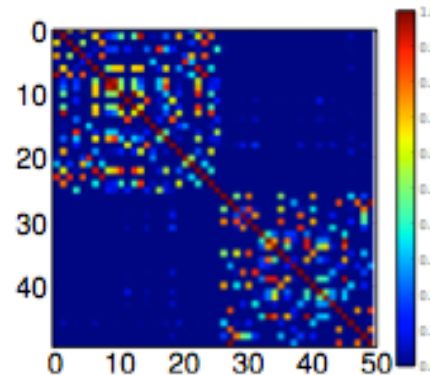
Example: Gaussian kernel

$$e^{-\frac{\|x - y\|^2}{2\sigma^2}}$$

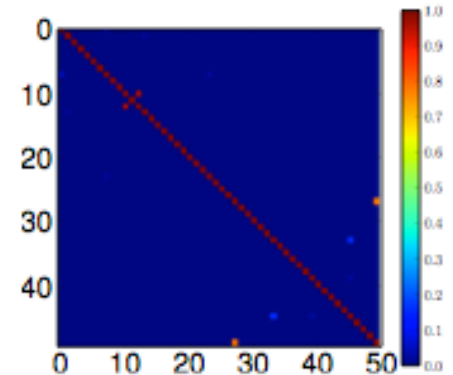
Blaschko @Philips 2013



$\sigma = 1$



$\sigma = 10$

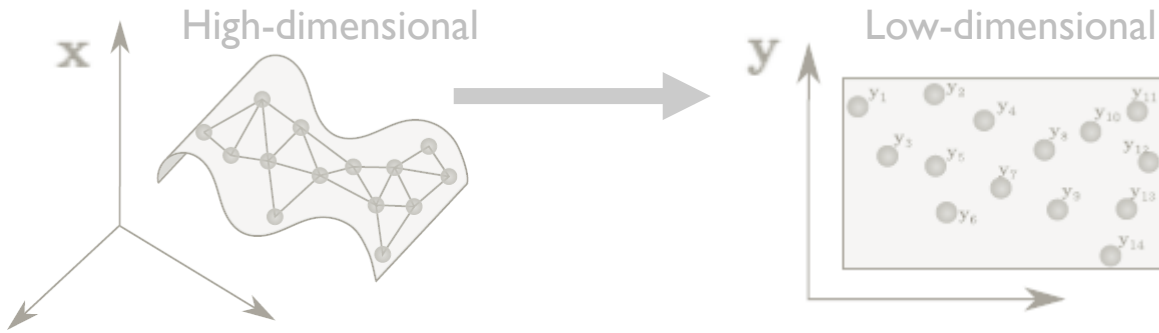


$\sigma = 1000$

Apprendre à représenter une population: manifold learning / graph embedding

Réduction de dimension:

$$f : \mathbf{x} \in \mathbb{R}^h \mapsto \mathbf{y} \in \mathbb{R}^l \quad h \gg l$$



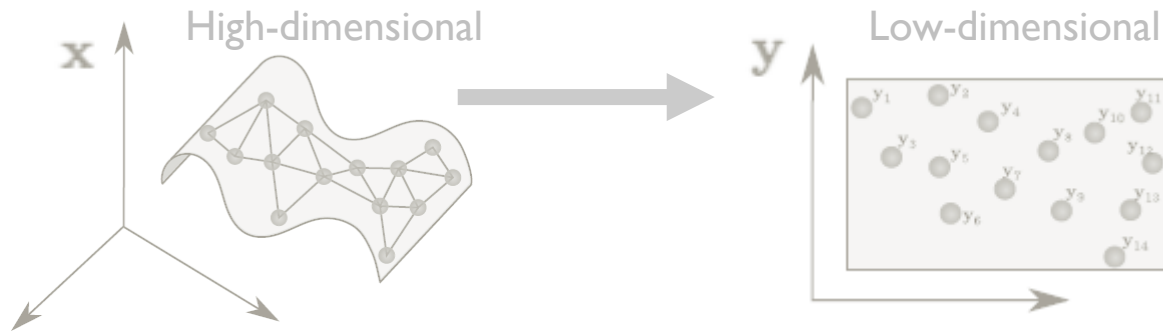
Method	Operator/Matrix	Preserved	Objective Function
PCA	Covariance matrix	Variance of the dataset / Euclidean distances between data points	$\mathbf{u}^T \Sigma \mathbf{u}$
Laplacian Eigenmaps	Graph Laplacian	Distances within the local neighbourhood of each data point	$\mathbf{u}^T L \mathbf{u}$
ISOMAP	Geodesic distance matrix	Geodesic distances between data points	$\mathbf{u}^T D_G \mathbf{u}$
LLE	Reconstruction weights	Reconstruction weights within the local neighbourhood of each data point	$\mathbf{u}^T W \mathbf{u}$

Atasoy & Mateus, MICCAI 2011

Apprendre à représenter une population: manifold learning / graph embedding

Réduction de dimension:

$$f : \mathbf{x} \in \mathbb{R}^h \mapsto \mathbf{y} \in \mathbb{R}^l \quad h \gg l$$



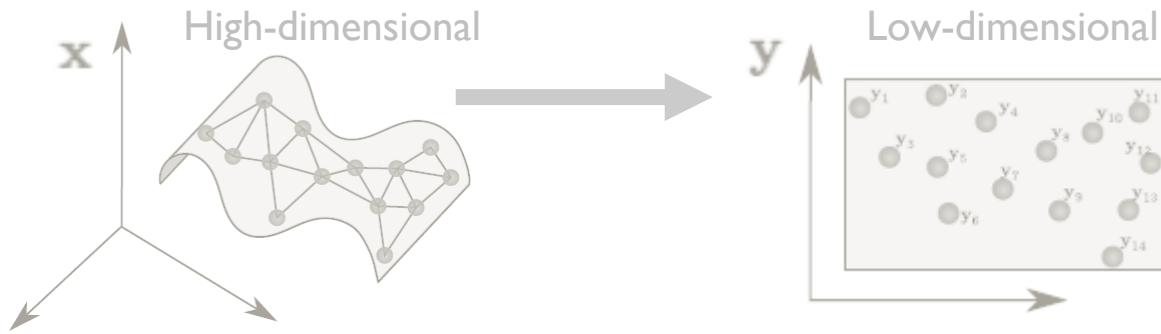
Revient à résoudre:

$$\hat{\mathbf{Y}} = \operatorname{argmin} \sum_{i,j} \|\mathbf{y}_i - \mathbf{y}_j\|^2 w_{ij} = \operatorname{argmin} \mathbf{Y}^t \mathbf{L} \mathbf{Y}$$

Yan et al. IEEE PAMI 2007

Apprendre à représenter une population: manifold learning / graph embedding

Réduction de dimension: $f : \mathbf{x} \in \mathbb{R}^h \mapsto \mathbf{y} \in \mathbb{R}^l$ $h \gg l$



Revient à résoudre: $\hat{\mathbf{Y}} = \operatorname{argmin} \sum_{i,j} \|\mathbf{y}_i - \mathbf{y}_j\|^2 w_{ij} = \operatorname{argmin} \mathbf{Y}^t \mathbf{L} \mathbf{Y}$

Yan et al. IEEE PAMI 2007

Sous contrainte:

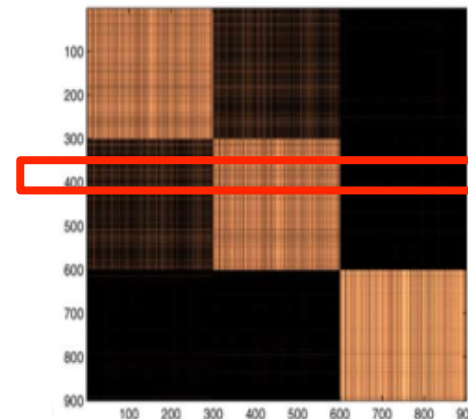
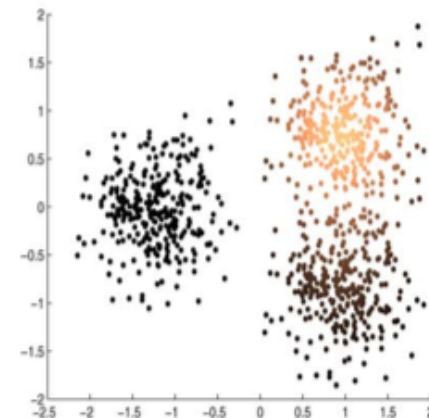
Unsupervised $\sum_i \|\mathbf{y}_i\|^2 d_{ii} = 1$

Supervised $\sum_{i,j} \|\mathbf{y}_i - \mathbf{y}_j\|^2 w'_{ij} = 1$

graph Laplacian
 $\mathbf{L} = \mathbf{D} - \mathbf{W}$

$$d_{ii} = \sum_j w_{ij}$$

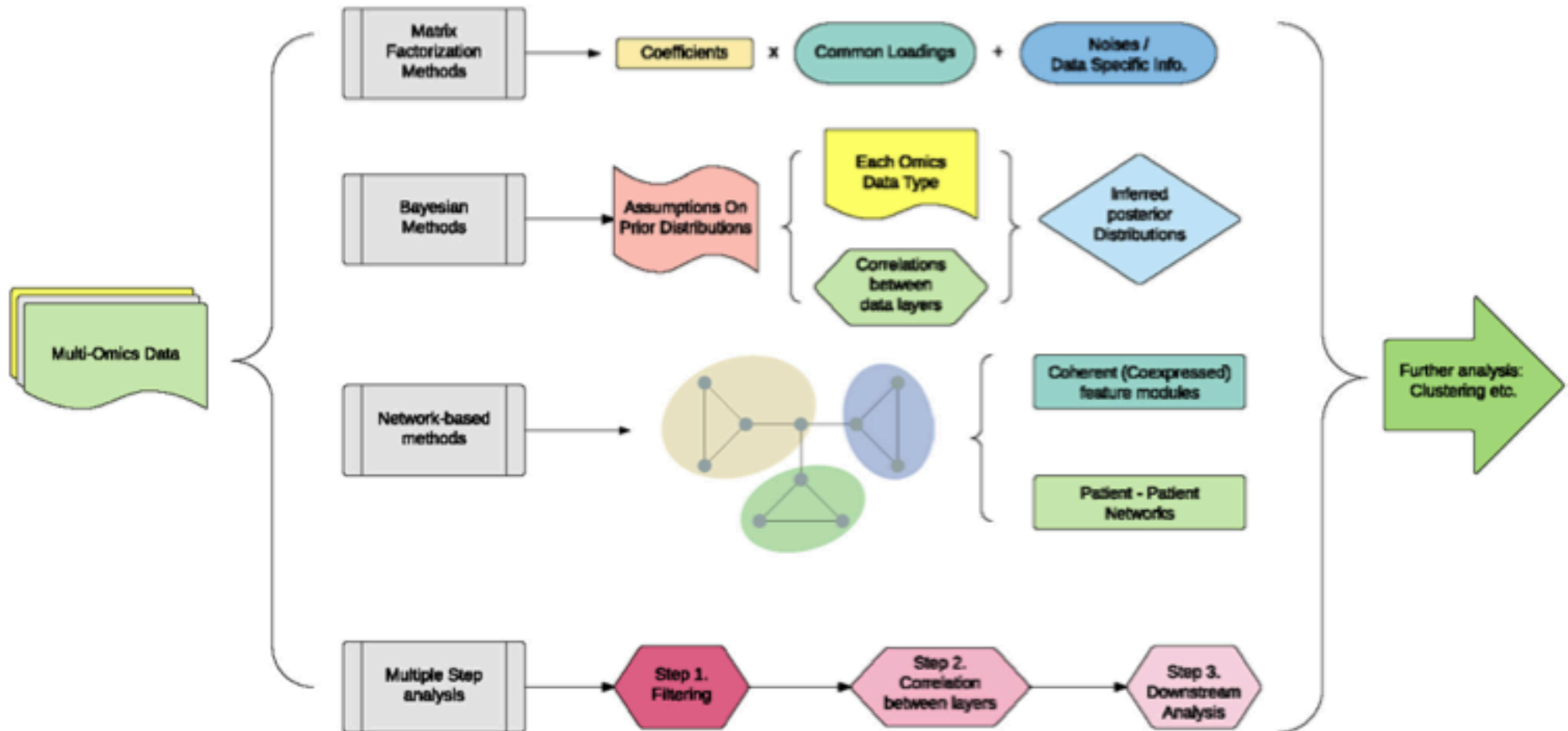
Coifman et al. ACHA, 2006



Mélange de descripteurs multiples

➤ Valable en haute dimension ?

Huang et al. *Front Genet* 2017



Mélange de descripteurs multiples

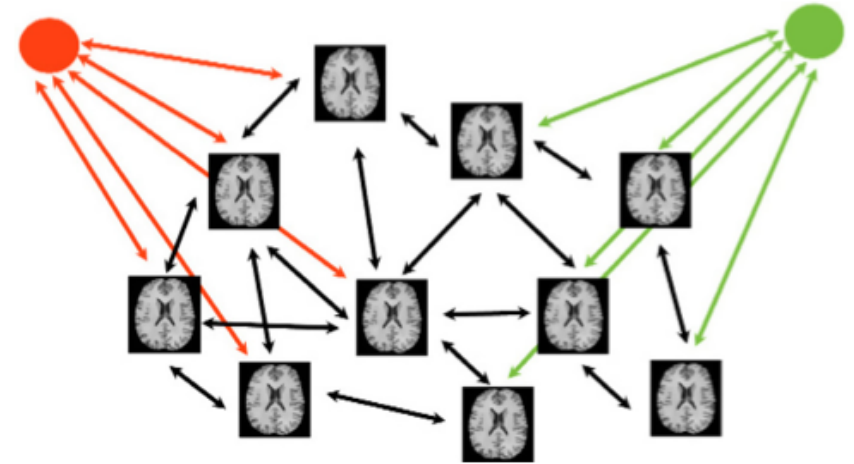
PhD of S. Sanchez @UPF Barcelona, ES (2018)

Comment mélanger des données **hétérogènes** ?

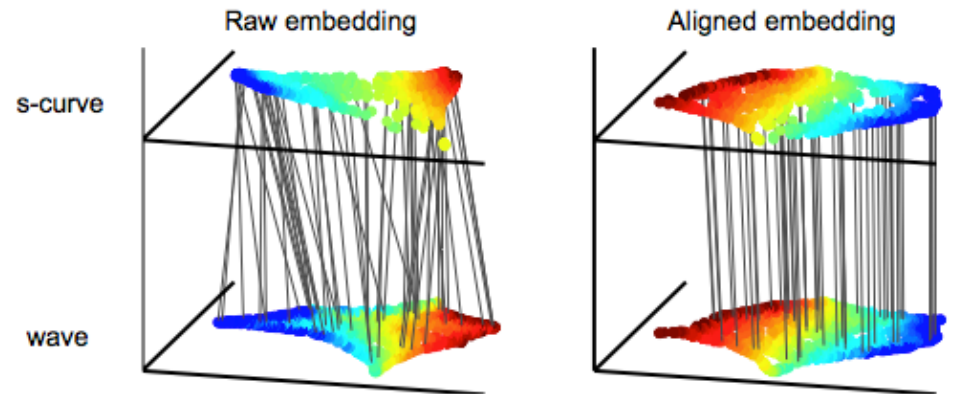


Normalisation
+ concaténation

Noyaux (kernels)



Wolz et al. *Med Image Anal* 2012



Ham et al. *AISTATS* 2005
 Valencia et al. *CIARP* 2011
 Lee et al. *Patt Recogn* 2016

Mélange de descripteurs multiples

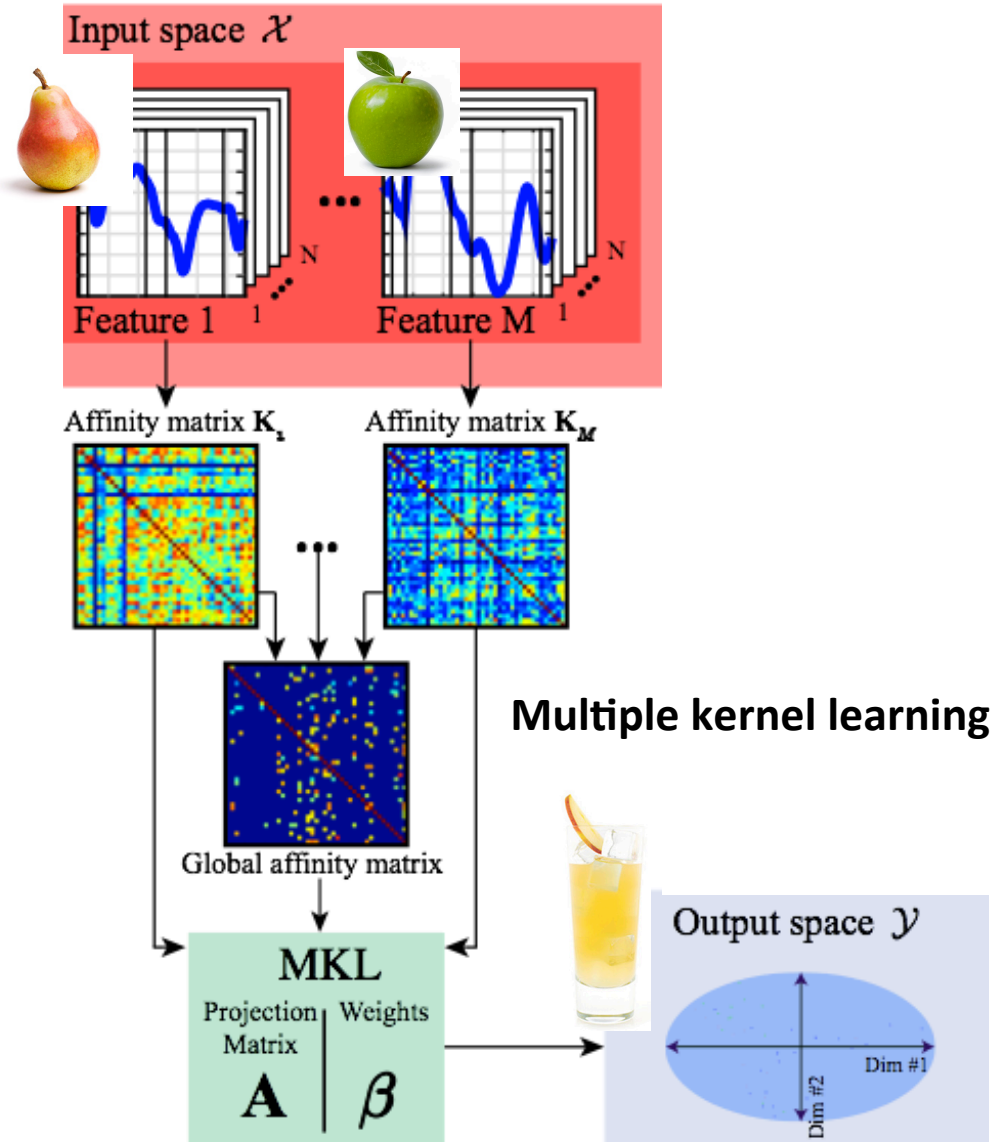
PhD of S. Sanchez @UPF Barcelona, ES (2018)

Comment mélanger des données hétérogènes ?



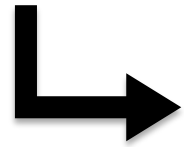
Normalisation + concaténation

Noyaux (kernels)



Mélange de descripteurs multiples

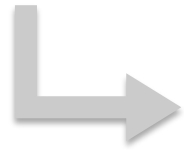
Linéaire, 1 descripteur: $\mathbf{y}_i = \mathbf{v}^t \mathbf{x}_i$



$$\hat{\mathbf{v}} = \underset{\mathbf{Y}^t \mathbf{B} \mathbf{Y} = \delta}{\operatorname{argmin}} \sum_{i,j} \|\mathbf{v}^t \mathbf{x}_i - \mathbf{v}^t \mathbf{x}_j\|^2 w_{ij} = \underset{\mathbf{Y}^t \mathbf{B} \mathbf{Y} = \delta}{\operatorname{argmin}} \mathbf{v}^t \mathbf{X} \mathbf{L} \mathbf{X}^t \mathbf{v}$$

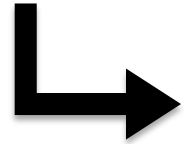
Mélange de descripteurs multiples

Linéaire, 1 descripteur: $\mathbf{y}_i = \mathbf{v}^t \mathbf{x}_i$



$$\hat{\mathbf{v}} = \underset{\mathbf{Y}^t \mathbf{B} \mathbf{Y} = \delta}{\operatorname{argmin}} \sum_{i,j} \|\mathbf{v}^t \mathbf{x}_i - \mathbf{v}^t \mathbf{x}_j\|^2 w_{ij} = \underset{\mathbf{Y}^t \mathbf{B} \mathbf{Y} = \delta}{\operatorname{argmin}} \mathbf{v}^t \mathbf{X} \mathbf{L} \mathbf{X}^t \mathbf{v}$$

Non-linéaire, multiples descripteurs:



$$\underset{\mathbf{A}, \beta}{\operatorname{argmin}} \sum_{i,j=1}^N \|\mathbf{A}^t \mathbf{K}^{(i)} \beta - \mathbf{A}^t \mathbf{K}^{(j)} \beta\|^2 w_{ij}$$

Fusion des descripteurs:
$$\mathbf{v}^t \Phi(\mathbf{x}_i) = \sum_{n=1}^N \sum_{m=1}^M \alpha_n \beta_m k_m(\mathbf{x}_n, \mathbf{x}_i) = \mathbf{A}^t \mathbf{K}^{(i)} \beta$$

Mélange de descripteurs multiples

$$\underset{\mathbf{A}, \beta}{\operatorname{argmin}} \sum_{i,j=1}^N \|\mathbf{A}^t \mathbf{K}^{(i)} \beta - \mathbf{A}^t \mathbf{K}^{(j)} \beta\|^2 w_{ij}$$

En pratique: optimisation alternée

1. Fix \mathbf{A} and optimize β

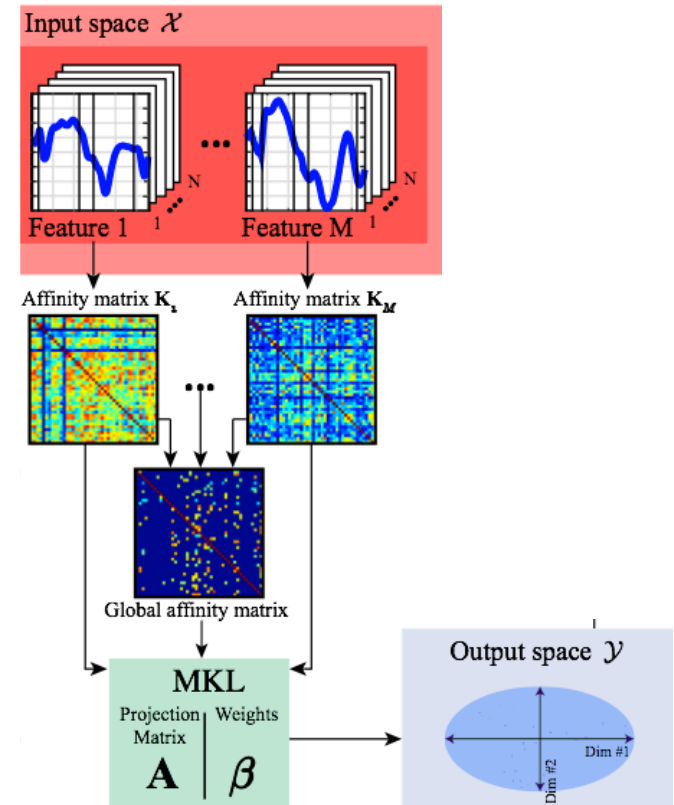
= solving a trace-ratio problem

2. Fix β and optimize \mathbf{A}

= not anymore a generalized eigenvalue problem due to $\beta_m \geq 0$

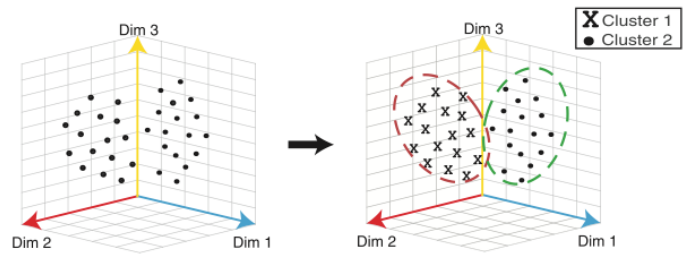
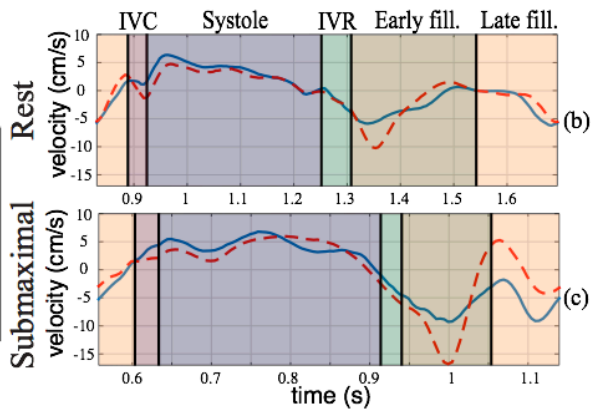
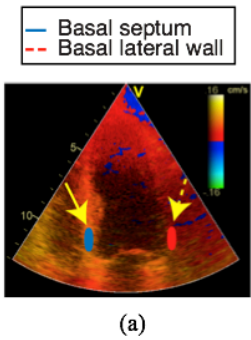
Problem: non-convex, hard to solve

> Can be relaxed and solved by semidefinite programming



Mélange de descripteurs multiples

= fraction d'éjection préservée (HFPEF)

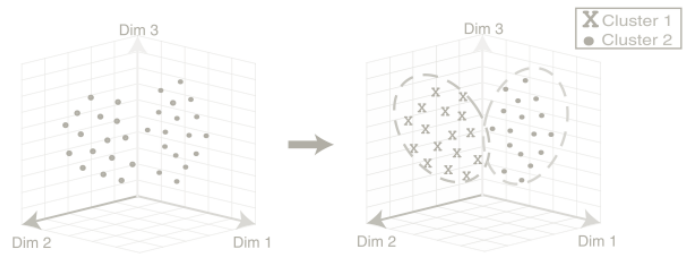
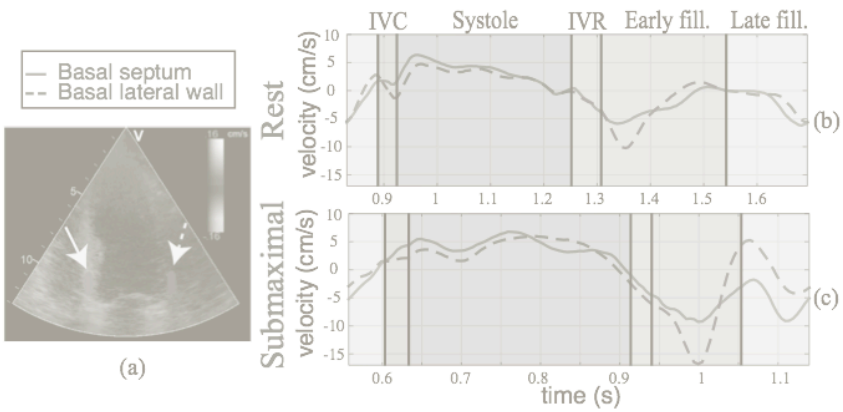


Continuum de normal à pathologique ?

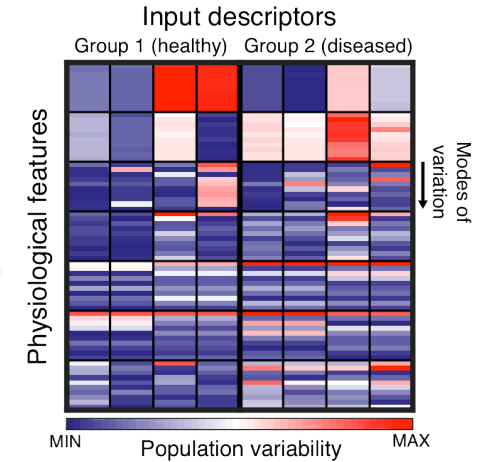
Pas de confiance dans les labels → non-supervisé ?

Mélange de descripteurs multiples

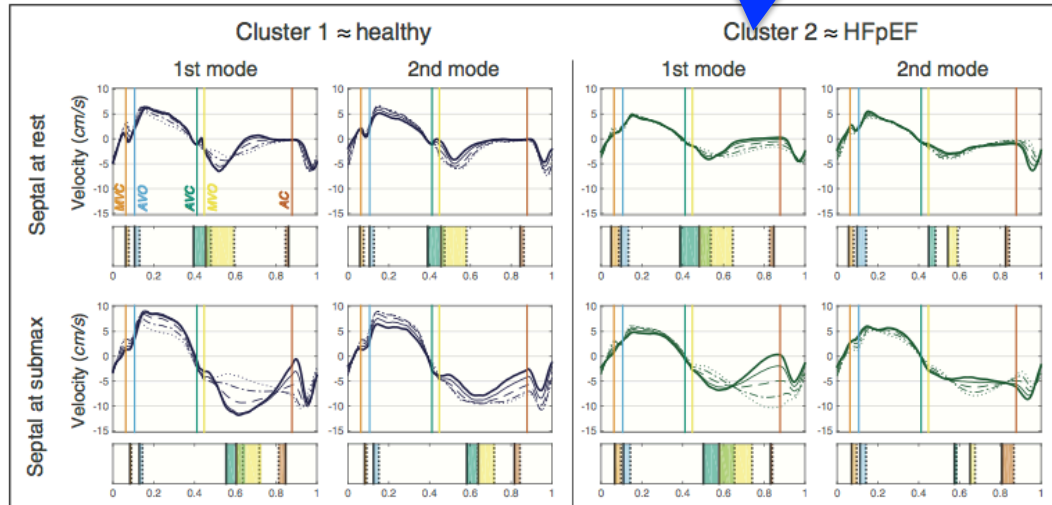
= fraction d'éjection préservée (HFPEF)



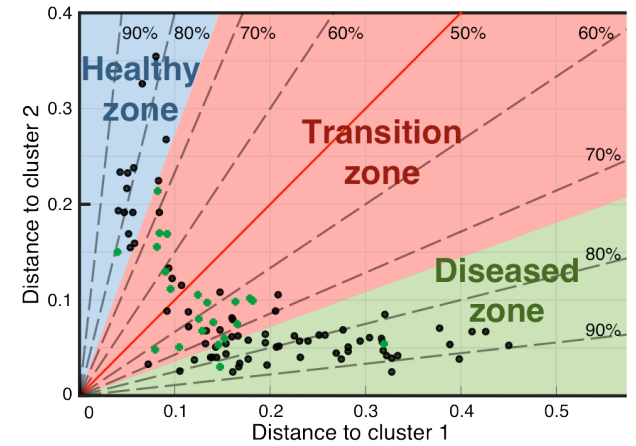
Phénotypage



Analyse de variabilité

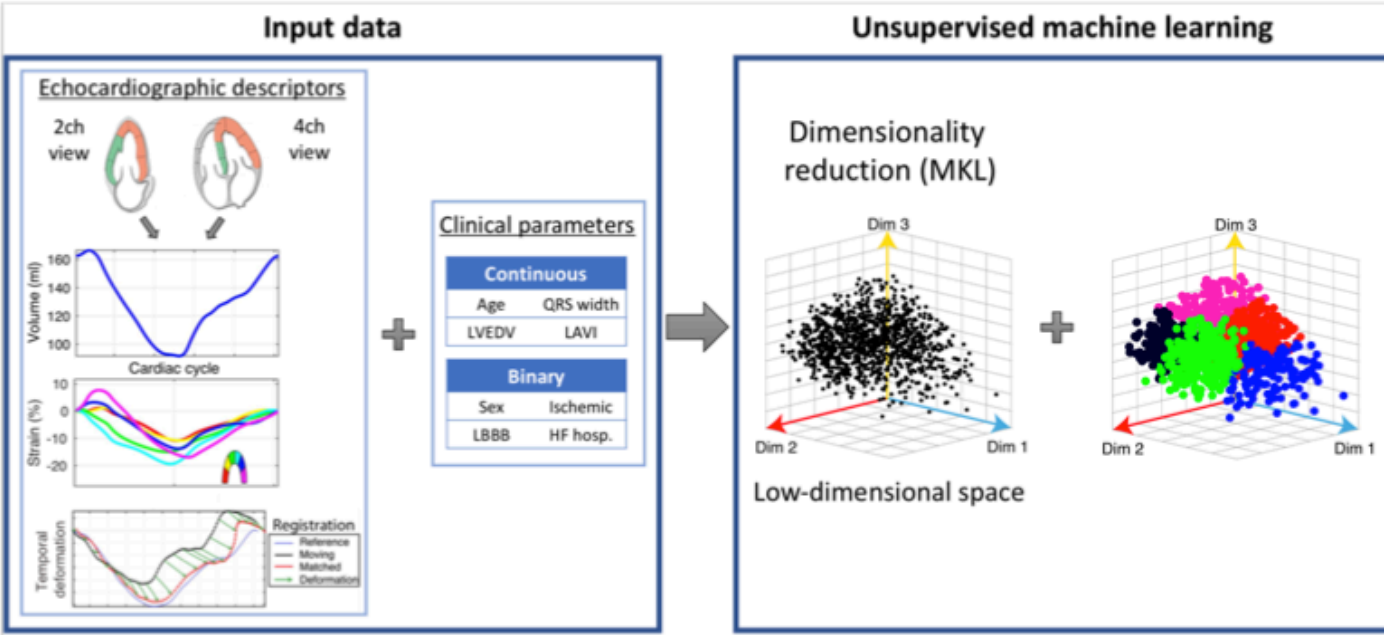


Stratification de risque



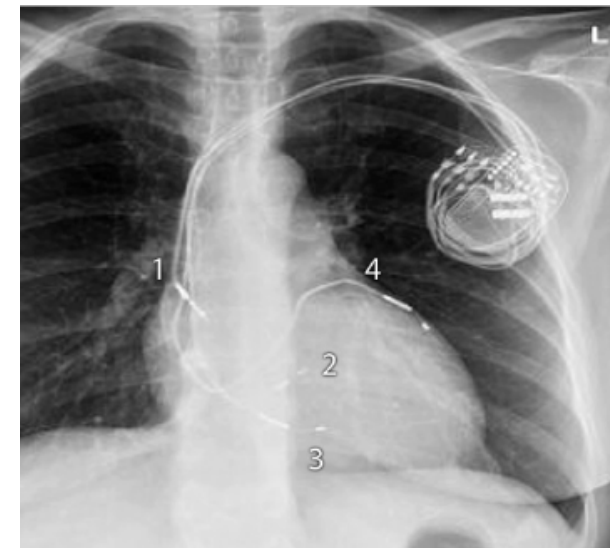
Mélange de descripteurs multiples

= resynchronisation cardiaque (CRT)



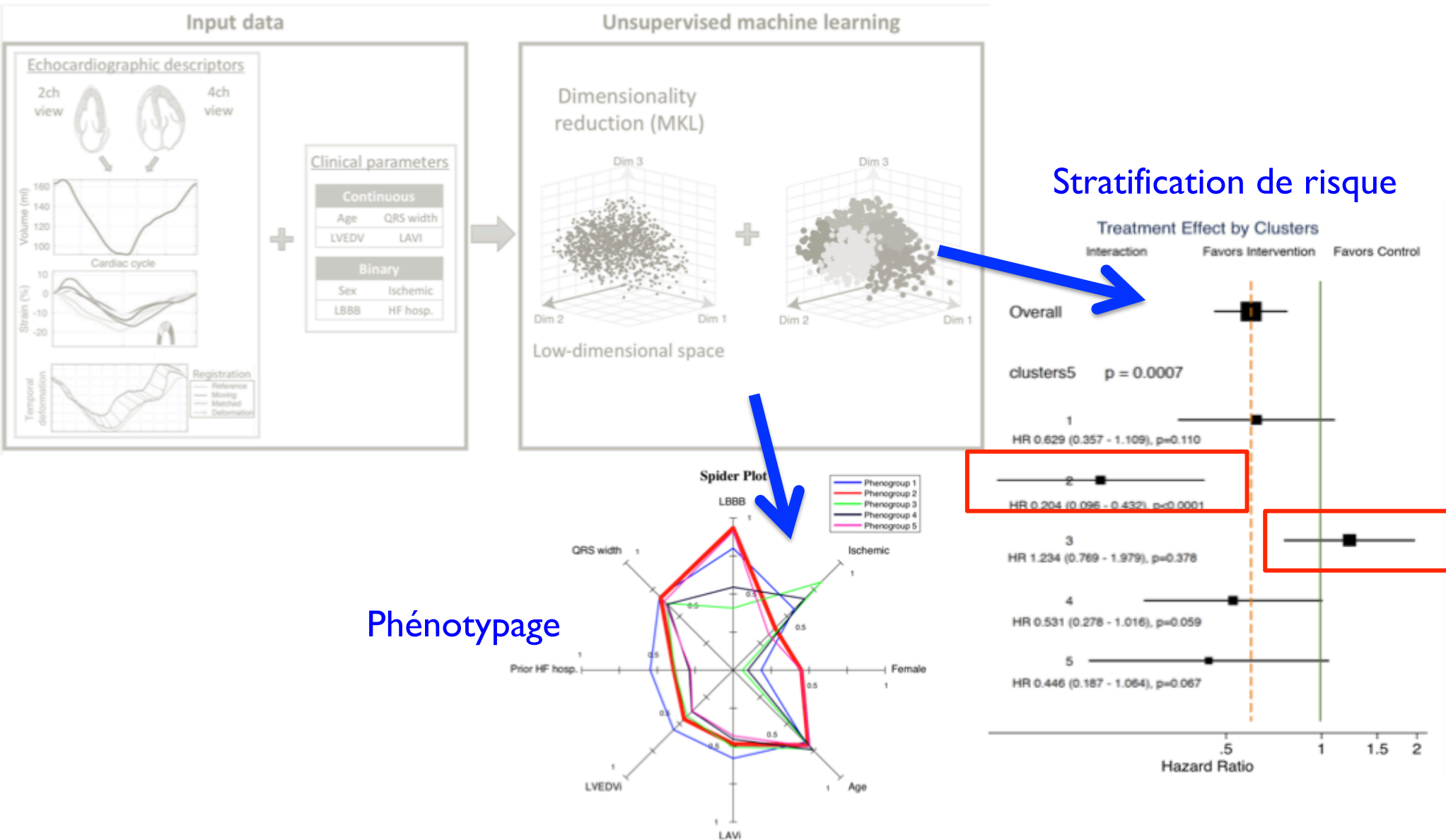
Compréhension de la réponse à la thérapie ?
(30% de non-répondeurs !!!)

Limites d'un raisonnement supervisé → non supervisé ?



Mélange de descripteurs multiples

= resynchronisation cardiaque (CRT)



Conclusions

- **Cadre unifié** pour représenter des populations
- Possibilité de mélange de données **hétérogènes et haute-dimension**
 - Sortie = espace faible dimension + poids de chaque descripteur



Conclusions

- **Cadre unifié** pour représenter des populations
- Possibilité de mélange de données **hétérogènes et haute-dimension**
 - Sortie = espace faible dimension + poids de chaque descripteur

Et après...

- **Interprétation** des résultats ?
- Cadre **physiologique** ?
- Tout **mélanger** à la fois ?
- **Structuration** des entrées ?

