Statistiques sur les variétés différentielles homogènes

Application aux primitives géométriques en traitement d'image 3D

Xavier Pennec

INRIA Sophia-Antipolis
Projet EPIDAURE
B.P. 93, 2004 Route des Lucioles
06 902 Sophia Antipolis Cedex, France

Email : xavier.pennec@sophia.inria.fr
http://www.inria.fr/epidaure/Epidaure-eng.html

Plan de l'exposé

- Cadre du problème
- Géométrie riemannienne
- Probabilités et statistiques
- Implantation et applications
  - Imagerie médicale
  - Biologie moléculaire
- Conclusion

Traitement d'image haut niveau

OBJETS : modèles géométriques structurés

Reconnaissance
Modélisation
Mise en correspondance
Recalage
Interprétation
Classification

Traitements haut niveau

SCÈNE : Primitives géométriques non structurées

Segmentation
Extraction de contours
Reconstruction
Filtrage
Rééchantillonage

Traitements bas niveau

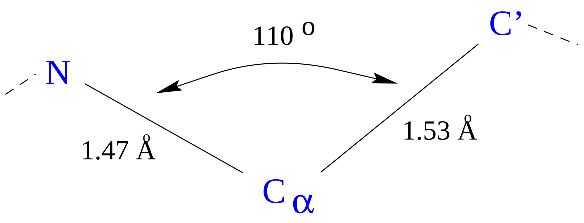
DONNÉES : Images 2D, 3D, spectres ...

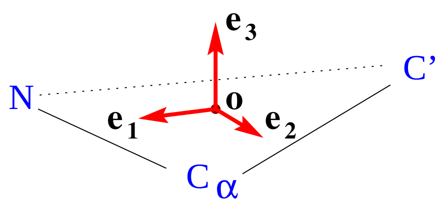
Biologie structurale
- *Cristallographie*

Imagerie
- *Vision 2D*
- *Données de profondeur (2D1/2)*
- *Imagerie 3D*

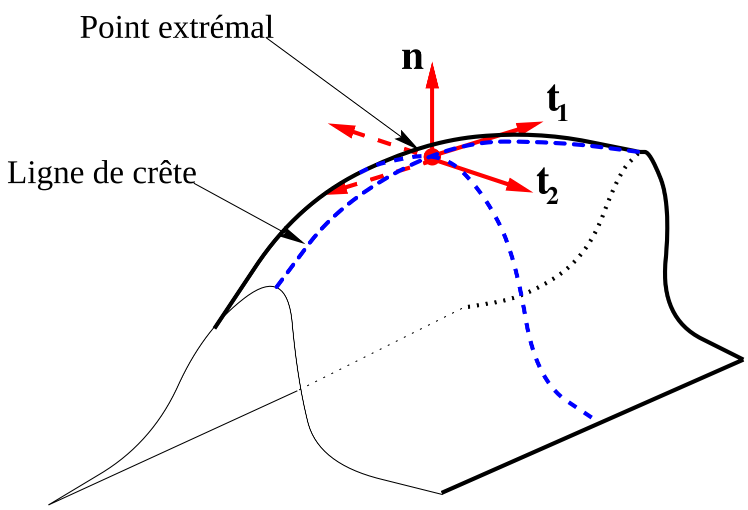
Exemples de primitives

Acides aminés : repères





Les points extrémaux : repères semi-orientés



Problématique

Pourquoi des primitives ?

- Réduction de complexité (matching)
- Sélectivité plus importante (matching)
- Précision plus importante (recalage)

Gestion fine de l'incertitude

- Reconnaissances fantômes
- Faux négatifs
- Précision

Méthodes génériques

- Briques algorithmiques

Probabilités et statistiques classiques

- **Mesure :** vecteur aléatoire $\boldsymbol{x}$ de densité $p_{\boldsymbol{x}}(y)$
- **Approximation :** $\boldsymbol{x} \sim (\overline{x}\ ,\ \Sigma_{\boldsymbol{x}\boldsymbol{x}})$
  - Moyenne : $\overline{x} = \mathbf{E}\left[\,\boldsymbol{x}\,\right] = \int y.p_{\boldsymbol{x}}(y).dy$
  - Covariance : $\Sigma_{\boldsymbol{x}\boldsymbol{x}} = \mathbf{E}\left[\,(\boldsymbol{x}-\overline{x}).(\boldsymbol{x}-\overline{x})^{\top}\,\right]$

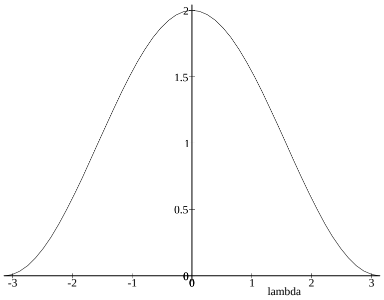
$$\Sigma_{\boldsymbol{x}\boldsymbol{x}} = \int (y-\overline{x}).(y-\overline{x})^{\top}.p_{\boldsymbol{x}}(y).dy$$

- **Propagation :**

$$\boldsymbol{y} = h(\boldsymbol{x}) \quad \sim \quad \left(h(\overline{x})\ ,\ \frac{\partial h}{\partial x}.\Sigma_{\boldsymbol{x}\boldsymbol{x}}.\frac{\partial h}{\partial x}^{\top} \right)$$

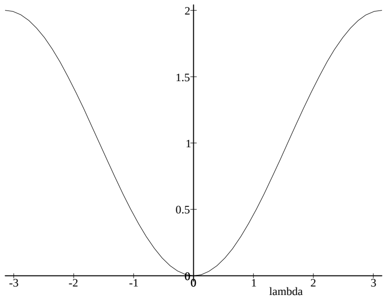
- **Modèle de bruit :** additif, gaussien ...
- **Distance Statistique :** Mahalanobis et χ^2

Espérances de droites vectorielles 2D



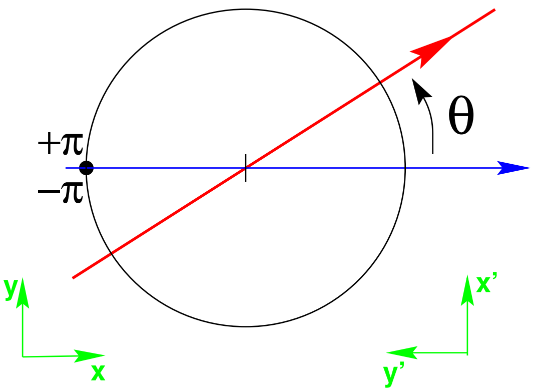
$$\rho_0(\theta) = 2\cos(\theta/2)^2$$

$$\overline{\theta}_0 = \mathbf{E}\left[\,\boldsymbol{\theta_0}\,\right] = 0$$

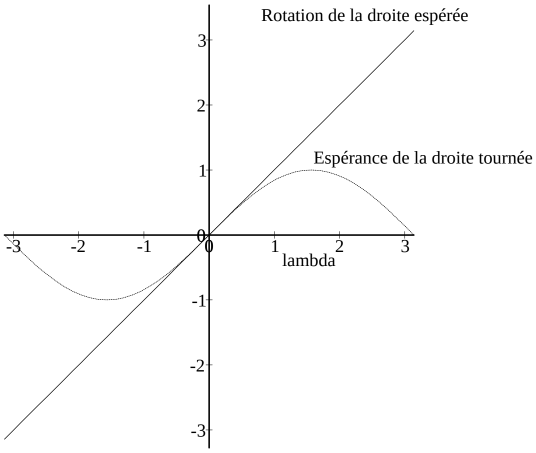


$$\rho_\pi(\theta) = 2\sin(\theta/2)^2$$

$$\overline{\theta_\pi} = 0 \text{ et } \underline{\theta_\pi} = \pi$$



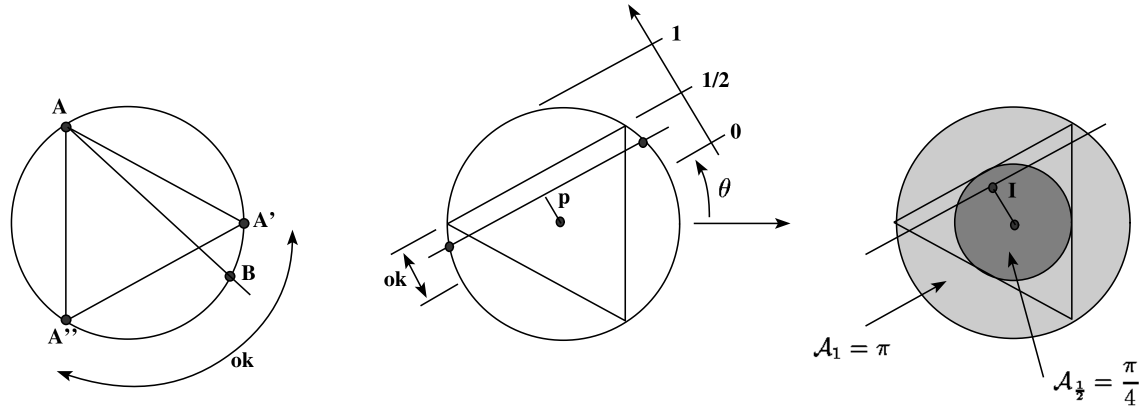
$$\theta \in]-\pi\,;\,\pi[$$



$$R_\lambda(\mathbf{E}\left[\,\theta\,\right]) \text{ et } \mathbf{E}\left[\,R_\lambda(\theta)\,\right]$$

Le paradoxe de Bertrand (1907)

Probabilité qu'une corde prise "au hasard" soit plus longue que le côté d'un triangle équilatéral inscrit ?



Méthode 1 $p = \frac{1}{3}$

Méthode 2 $p = \frac{1}{2}$

Méthode 3 $p = \frac{1}{4}$

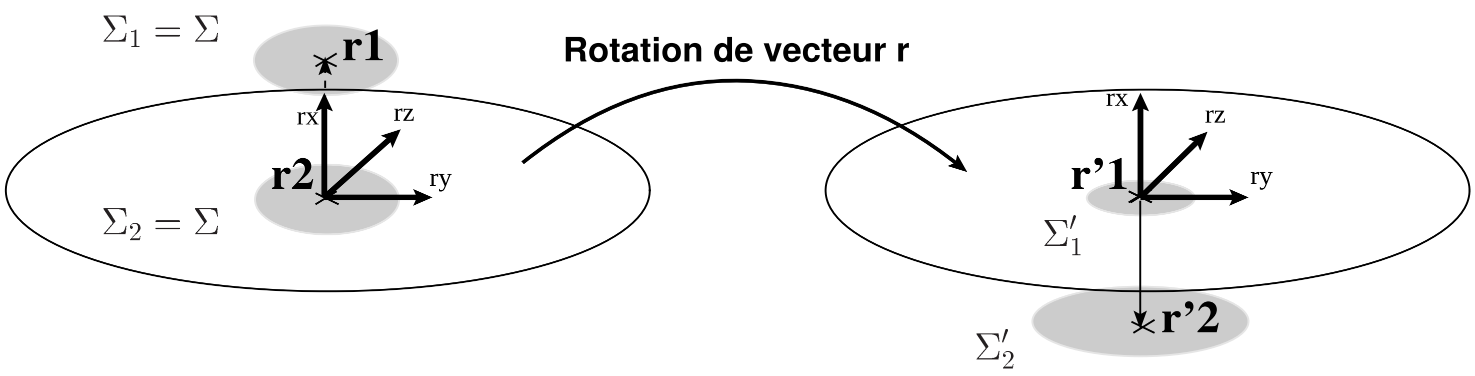
Densités correspondantes

$$d\sigma_1 = \frac{dp.d\theta}{2\pi\sqrt{1-p^2}} \qquad d\sigma_2 = \frac{dp.d\theta}{2\pi} \qquad d\sigma_3 = p.\frac{dp.d\theta}{\pi}$$

Choix : hypothèse d'invariance

Rotation et translation : $p = \frac{1}{2}$

Bruit sur les rotations



$r1 = (\frac{\pi}{2}, 0, 0) \qquad r2 = (0, 0, 0)$ $\qquad\qquad r1 = (0, 0, 0) \qquad r2 = (-\frac{\pi}{2}, 0, 0)$

$$r = (-\frac{\pi}{2}, 0, 0)$$

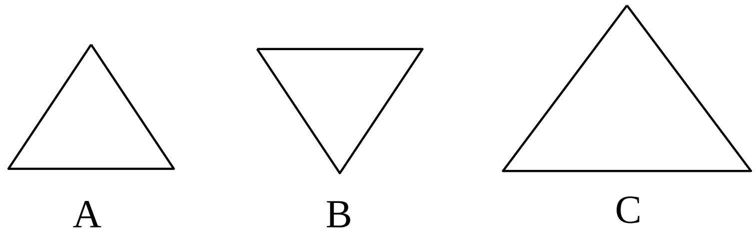
$$\Sigma_1 = \Sigma_2 = \sigma^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \Sigma \qquad\qquad \Sigma_1 = \frac{8}{\pi^2}.\Sigma < \Sigma_2 = \frac{\pi^2}{8}.\Sigma$$

Géométrie riemannienne

- Le cadre géométrique
- Mesure invariante
- Métrique riemanienne invariante
- Carte exponentielle

Le Cadre géométrique

Primitives géométriques : variété différentielle $\mathcal{M}$



- Les trois triangles sont similaires.
- A et B seulement sont congruents.
- Ils sont tous différents en translation.

Transformations : groupe de Lie $\mathcal{G}$

Opérations géométriques atomiques

- Composition : $f = f_1 \circ f_2$
- Inversion : $f^{(-1)}$
- Action : $y = f \star x$

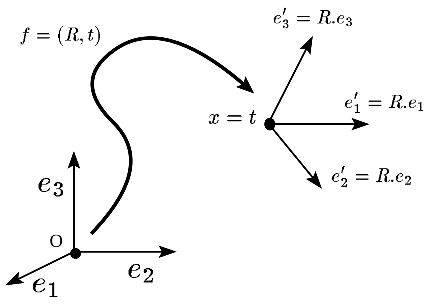
Cartes et atlas : représentation minimale

Le Cadre géométrique (suite)

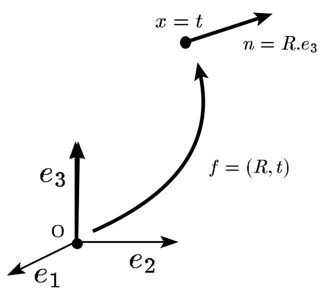
Variété homogène

- Origine : $\mathrm{o} \in \mathcal{M}$
- Groupe d'isotropie : $\mathcal{H} = \{\mathrm{f} \in \mathcal{G} \,/\, \mathrm{f} \star \mathrm{o} = \mathrm{o}\}$
- Variété homogène : $\mathcal{G} \star \mathrm{o} = \mathcal{M}$ donc $\mathcal{M} = \mathcal{G} \backslash \mathcal{H}$
- Cosets : $\mathcal{F}_\mathrm{x} = \{\mathrm{f} \in \mathcal{G} \,/\, \mathrm{f} \star \mathrm{o} = \mathrm{x}\}$
- Fonction de placement : $\mathrm{f}_\mathrm{x} \in \mathcal{F}_\mathrm{x}$

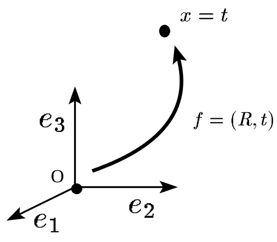
Exemples



Repères



Points orientés



Points

Mesure invariante

Mesure invariante sur $\mathcal{M}$

$$[\quad d(x_o + x) = dx \quad]$$

$$d\mathcal{M}(\mathrm{f}_0 \star x) = d\mathcal{M}(x)$$

Condition d'existence

$$\forall h \in \mathcal{H}, \qquad \left| \left. \frac{\partial (h \star e)}{\partial e} \right|_{e=o} \right| = 1$$

Représentation de la mesure

$$d\mathcal{M}(x) = \frac{dx}{|J(f_x)|} \qquad \text{avec} \qquad J(f_x) = \left. \frac{\partial (f_x \star e)}{\partial e} \right|_{e=o}$$

Mesure invariante sur les rotations $r = \theta.n$

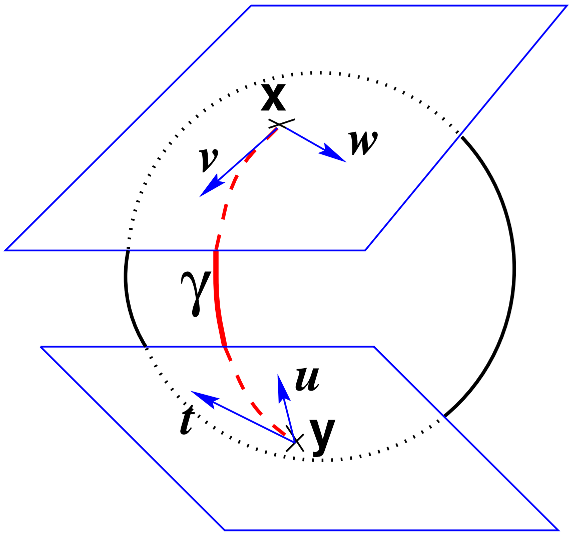
$$d\mathcal{M}\,(r) = \frac{4\sin^2(\theta/2)}{\theta^2}.dr$$

Géométrie riemannienne

- Métrique riemannienne :

Collection de produits scalaires sur les espaces tangents.

$$\langle v \mid w \rangle_x = v^\top . Q(x) . w$$



- Longueur d'une courbe : $\mathcal{L}(\gamma) = \int_a^b \|\dot{\gamma}(t)\|_{\gamma(t)} . dt$

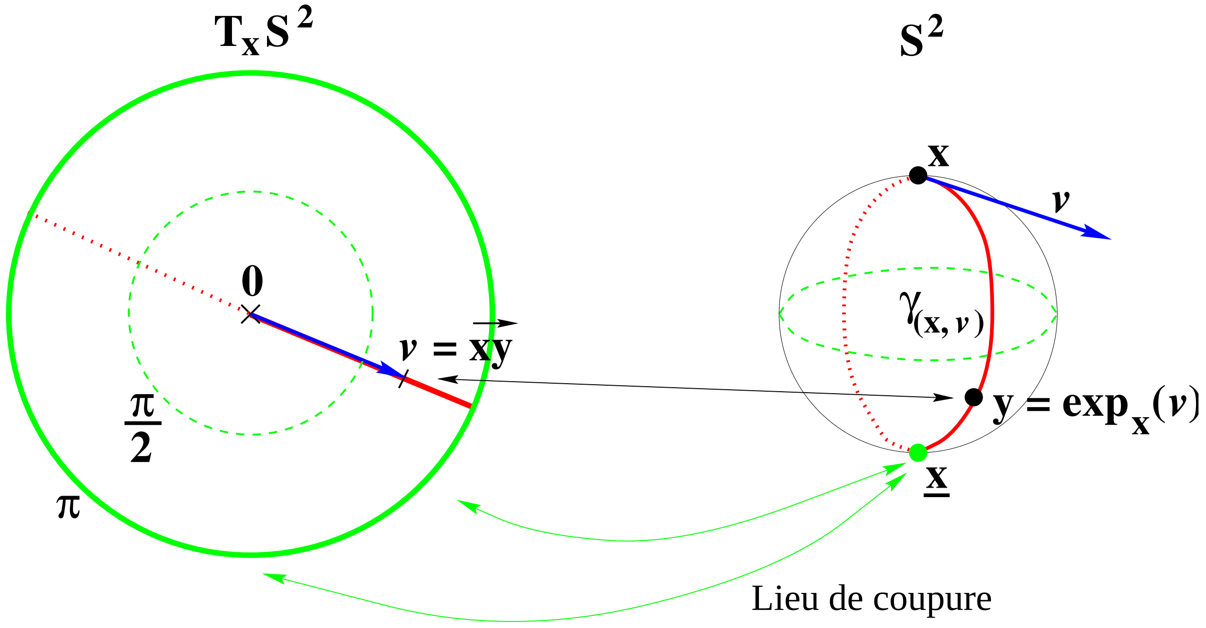
- Distance (variété connexe) : $\mathrm{dist}(\mathrm{x}, \mathrm{y}) = \inf_\gamma \mathcal{L}(\gamma)$

- Il existe une et une seule **géodésique** $\gamma_{(x,v)}(t)$

- Variété connexe géodésiquement complète

Géométrie riemannienne (suite)

Carte exponentielle : $\exp_x(v) = \gamma_{(x,v)}(1)$



Lieu de coupure : domaine maximal

Métrique invariante sur le groupe

Translation de l'identité : $\mathrm{f} = \mathrm{f} \circ \mathrm{Id} = \mathrm{Id} \circ \mathrm{f}$

$$J_L(f) = \left.\frac{\partial (f \circ e)}{\partial e}\right|_{e=Id}$$

Métrique invariante à gauche :

$$Q(f) = J_L(f)^{-\top}.Q.J_L(f)^{(-1)}$$

Carte principale : $\vec{\mathrm{f}} \longleftrightarrow \mathrm{f}$

Cartes exponentielles : $\overrightarrow{\mathrm{gf}} = J_L(\vec{\mathrm{g}}).(\vec{\mathrm{g}}^{(-1)} \circ \vec{\mathrm{f}})$

Distance (invariante à gauche) :

$$\mathsf{dist}(\mathrm{f}, \mathrm{g})^2 = \| \overrightarrow{\mathrm{fg}} \|_{\mathrm{f}}^2 = (\vec{\mathrm{f}}^{(-1)} \circ \vec{\mathrm{g}})^\top.Q.(\vec{\mathrm{f}}^{(-1)} \circ \vec{\mathrm{g}}) = \|\vec{\mathrm{f}}^{(-1)} \circ \vec{\mathrm{g}}\|_{\mathrm{Id}}^2$$

Métrique invariante sur la variété

Translation de l'origine : $\mathrm{f_x} \in \mathcal{F}_\mathrm{x}$

$$\mathrm{x} = \mathrm{f_x} \star \mathrm{o} \qquad \text{et} \qquad J(f_x) = \left. \frac{\partial (f_x \star e)}{\partial e} \right|_{e=\mathrm{o}}$$

Métrique invariante à gauche :

$$Q(x) = J(\mathrm{f}_x)^{-\top}.Q.J(\mathrm{f}_x)^{\text{(-1)}}$$

Condition d'existence :

$$\forall \mathrm{h} \in \mathcal{H} \qquad J(\mathrm{h})^\top.Q.J(\mathrm{h}) = Q$$

Cartes exponentielles : $\overrightarrow{\mathrm{xy}} = J(\mathrm{f}_{\vec{\mathrm{x}}}).(\mathrm{f}_{\vec{\mathrm{x}}}^{\text{(-1)}} \star \vec{\mathrm{y}})$

Distance (invariante) :

$$\mathrm{dist}(\mathrm{x},\mathrm{y})^2 = \|\overrightarrow{\mathrm{xy}}\|^2_\mathrm{x} = (\mathrm{f}_{\vec{\mathrm{x}}}^{\text{(-1)}} \circ \vec{\mathrm{y}})^\top.Q.(\mathrm{f}_{\vec{\mathrm{x}}}^{\text{(-1)}} \circ \vec{\mathrm{y}}) = \|\mathrm{f}_{\vec{\mathrm{x}}}^{\text{(-1)}} \circ \vec{\mathrm{y}}\|^2_o$$

Probabilités sur les variétés

- Densité de probabilité
- Moyenne et espérance de Fréchet
- Matrice de covariance
- Modèles de bruit
- Autres problèmes statistiques

Densité de probabilité

$$\forall \mathcal{X},\ P(\mathbf{x} \in \mathcal{X}) = \int_{\mathcal{X}} p_{\mathbf{x}}(\mathrm{y}).d\mathcal{M}(\mathrm{y}) \quad \text{et} \quad P(\mathcal{M}) = 1$$

Densité dans la représentation : $\rho_{\boldsymbol{x}}(\mathrm{y}) = \dfrac{p_{\boldsymbol{x}}(y)}{|J(f_y)|}$

Action sur une primitive aléatoire

$$p_{(\mathrm{f}\star\mathbf{x})}(\mathrm{y}) = p_{\mathbf{x}}(\mathrm{f}^{(\text{-}1)} \star \mathrm{y})$$

$$\left[\qquad p_{(x_o+\boldsymbol{x})}(y) = p_{\boldsymbol{x}}(-x_o + y) \qquad \right]$$

Deux transformations aléatoires

$$p_{(\mathbf{f_1} \circ \mathbf{f_2})}(\mathrm{f}) = \int_{\mathcal{G}} p_{\mathbf{f_1}}(\mathrm{g}).p_{\mathbf{f_2}}(\mathrm{g}^{(\text{-}1)} \circ \mathrm{f}).d_L\mathcal{G}(\mathrm{g})$$

$$\left[p_{(\boldsymbol{x}+\boldsymbol{y})}(z) = p_{\boldsymbol{x}} \circledast p_{\boldsymbol{y}}(z) = \int_{\mathbb{R}^n} p_{\boldsymbol{x}}(t).p_{\boldsymbol{y}}(z-t).dt \right]$$

Espérance d'une observable $\varphi : \mathcal{M} \mapsto \mathbb{R}$

$$\mathbf{E}\left[\, \varphi(\boldsymbol{x}) \,\right] = \int_{\mathcal{M}} \varphi(y).p_{\boldsymbol{x}}(y).d\mathcal{M}(y)$$

Espérance de Fréchet (1944)

Variance par rapport à une primitive

$$\mathbf{E}\left[\ \mathrm{dist}(\mathrm{y}, \mathbf{x})^2\ \right] = \int_{\mathcal{M}} \mathrm{dist}(\mathrm{y}, \mathrm{z})^2 . p_{\mathbf{x}}(\mathrm{z}) . d\mathcal{M}(\mathrm{z})$$

$$\mathbf{E}\left[\ \mathrm{dist}(\mathrm{y}, \{\mathrm{x}_i\})^2\ \right] = \frac{1}{n} \sum_i \mathrm{dist}(\mathrm{y}, \mathrm{x}_i)^2$$

Primitive espérée ou moyenne

$$\mathbb{E}\left[\, \mathbf{x} \,\right] = \arg\min_{\mathrm{y} \in \mathcal{M}} \left(\mathbf{E}\left[\ \mathrm{dist}(\mathrm{y}, \mathbf{x})^2\ \right] \right)$$

Primitives centrales

$$\mathbb{E}^{\alpha}\left[\, \mathbf{x} \,\right] = \arg\min_{\mathrm{y} \in \mathcal{M}} \left(\mathbf{E}\left[\ \mathrm{dist}(\mathrm{y}, \mathbf{x})^{\alpha}\ \right] \right)^{1/\alpha}$$

Existence et unicité : Karcher et Kendall.

Espérance de Fréchet - Propriétés

Invariance

$$\mathbb{E}[\, \mathrm{f} \star \mathbf{x} \,] = \mathrm{f} \star \mathbb{E}[\, \mathbf{x} \,]$$

Caractérisation

$$\mathbf{E}\left[\, \overrightarrow{\bar{\mathrm{x}}\mathbf{x}} \,\right] = \int_{\mathcal{M}} \overrightarrow{\bar{\mathrm{x}}\mathrm{z}} \,.p_{\mathbf{x}}(\mathrm{z}).d\mathcal{M}(\mathrm{z}) = 0$$

$$\bar{\mathrm{x}} \in \mathbb{E}[\, \mathbf{x} \,] \quad \Longrightarrow \quad \mathbf{E}\left[\, \vec{\mathrm{f}}_{\bar{\mathrm{x}}}^{(-1)} \circ \vec{\mathbf{x}} \,\right] = 0$$

Cas des points rigides : espérance classique.

$$\overline{x} \in \mathbb{E}[\, \boldsymbol{x} \,] \quad \Longrightarrow \quad \mathbf{E}[\, -\overline{x} + \boldsymbol{x} \,] = 0$$

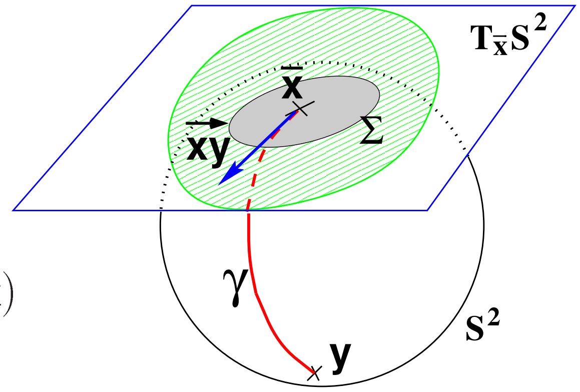
Algorithme d'obtention

$$\bar{\mathrm{x}}_{t+1} = \mathrm{f}_{\bar{\mathrm{x}}_t} \star \left(\frac{1}{n} \sum_i \left(\mathrm{f}_{\bar{\mathrm{x}}_t}^{(-1)} \star \vec{\mathrm{x}}_i \right) \right)$$

Matrice de covariance

$$\Sigma_{\mathbf{xx}} = \mathbf{E}\left[\, \overrightarrow{\bar{\mathrm{x}}\mathbf{x}}.\overrightarrow{\bar{\mathrm{x}}\mathbf{x}}^{\top} \,\right]$$

$$\Sigma_{\mathbf{xx}} = \int_{\mathcal{M}} (\overrightarrow{\bar{\mathrm{x}}\mathrm{x}}).(\overrightarrow{\bar{\mathrm{x}}\mathrm{x}})^{\top}.p_{\mathbf{x}}(\mathrm{x}).d\mathcal{M}(\mathrm{x})$$



Utilisation de la carte principale :

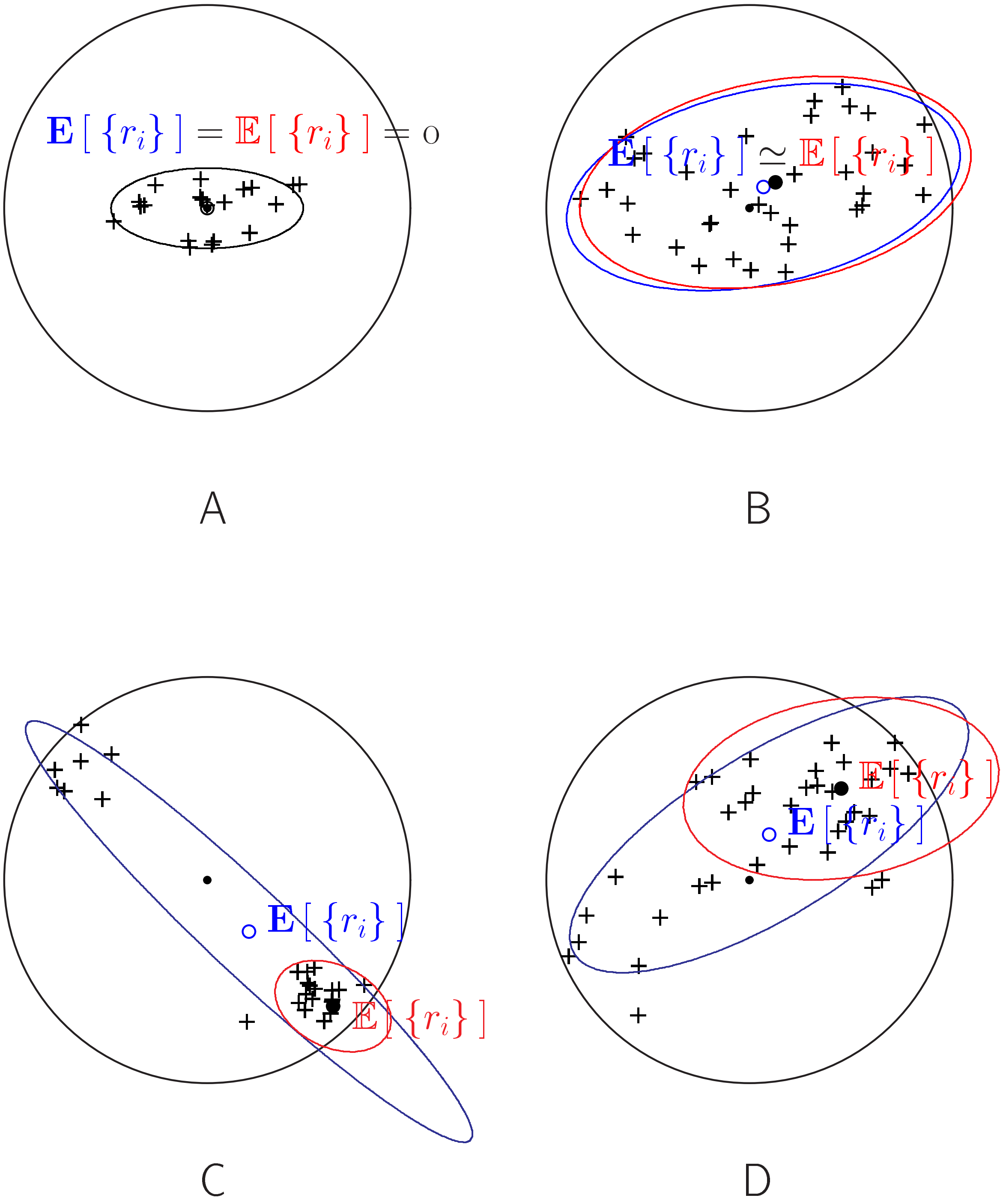
$$\mathbf{x} = f_{\bar{\mathrm{x}}} \star \mathbf{e} \quad \Longleftrightarrow \quad \mathbf{e} = f_{\bar{\mathrm{x}}}^{(-1)} \star \mathbf{x}$$

$$\Sigma_{\mathbf{ee}} = \mathbf{E}\left[\, \vec{\mathbf{e}}.\vec{\mathbf{e}}^{\top} \,\right] = \int_{\mathcal{D}} \vec{\mathrm{y}}.\vec{\mathrm{y}}^{\top}.\rho_{\vec{\mathbf{e}}}(y).dy$$

$$\Sigma_{\mathbf{xx}} = J(f_{\bar{x}}).\Sigma_{\mathbf{ee}}.J(f_{\bar{x}})^{\top}$$

Approximation pratique

$$\mathbf{x} \sim (\bar{\mathrm{x}}, \Sigma_{\mathbf{xx}})$$


$\mathbf{E}[\{r_i\}] = \mathbb{E}[\{r_i\}] = 0$
A
$\mathbf{E}[\{r_i\}] \simeq \mathbb{E}[\{r_i\}]$
B
$\mathbf{E}[\{r_i\}]$
$\mathbb{E}[\{r_i\}]$
C
$\mathbb{E}[\{r_i\}]$
$\mathbf{E}[\{r_i\}]$
D

Problèmes statistiques

Modèles de bruit

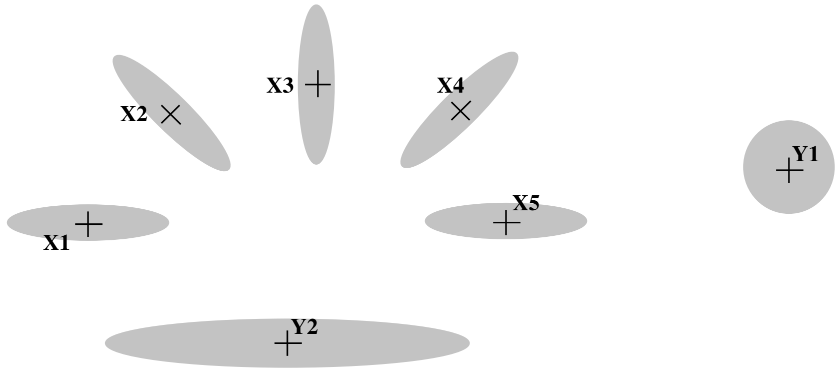
- Bruit homogène
- Bruit isotrope

Information

Distribution « gaussienne »

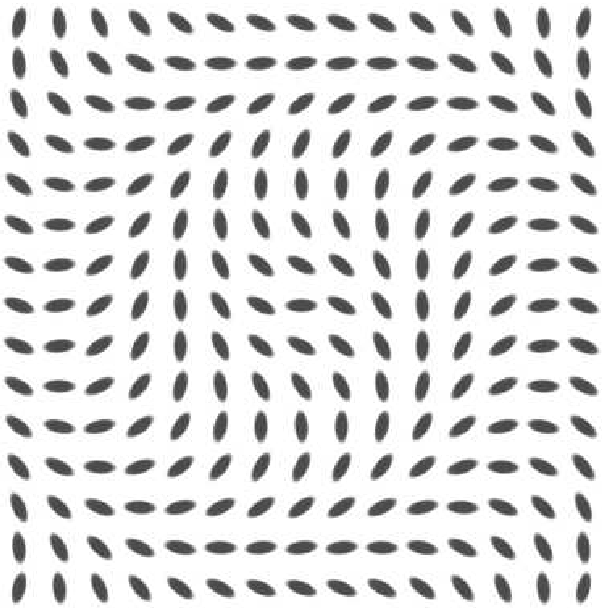
Distance de Mahalanobis

Modèles de bruit



Bruit homogène :

$$\exists f \in \mathcal{G} \;/\; \mathbf{x}(y) = f \star \mathbf{x}(z)$$



Bruit isotrope

$$\forall h \in \mathcal{H}, \;\; h \star \mathbf{e} = \mathbf{e}$$

Autres problèmes statistiques

Information

$$\mathbf{I}\,[\,\mathbf{x}\,] = -\mathbf{E}\,[\,\log(p_{\mathbf{x}}(\mathbf{x}))\,]$$

$$\mathbf{I}\,[\,\mathbf{x}\,] = -\int_{\mathcal{M}} \log(p_{\mathbf{x}}(\mathrm{x})).p_{\mathbf{x}}(\mathrm{x}).d\mathcal{M}(\mathrm{x})$$

Distribution « gaussienne »

$\Longrightarrow$ minimisation de l'information sachant

$$\mathbb{E}\,[\,\mathbf{x}\,] = \bar{\mathrm{x}} \qquad \text{et} \qquad \Sigma_{\mathbf{xx}} = \Sigma$$

Distance de Mahalanobis

$$\mu^2(\mathbf{x}, \mathrm{y}) = \overrightarrow{\bar{\mathrm{x}}\mathrm{y}}^{\top}.\Sigma_{\mathbf{xx}}^{(\text{-}1)}.\overrightarrow{\bar{\mathrm{x}}\mathrm{y}}$$

$$\mu^2(\mathbf{x}, \mathbf{y}) = \min_{\mathrm{z}} \left(\mu^2(\mathbf{x}, \mathrm{z}) + \mu^2(\mathbf{y}, \mathrm{z})\right)$$

Gaussienne : $\mathbf{E}\left[\,\mu^2(\mathbf{x}, \bar{x})\,\right] = n$

Implantation et applications

- Implantation « orientée objet »
- Recalage et validation
- Imagerie médicale
- Biologie moléculaire

Implantation « orienté objet »

Pour chaque type de primitive

- Déterminer les géodésiques et la carte principale
- Implanter les opérations canoniques dans cette représentation et leurs différentielles.
  - Groupe : composition, Inversion.
  - Variété : action, fonction de placement.

Transformations probabilistes : $\mathbf{f} \sim (\vec{\bar{f}}, \Sigma_{\mathbf{ff}})$

- Composition : $\mathbf{f} \circ \mathbf{g}$,
- Inversion : $\mathbf{f}^{(\text{-}1)}$,

Primitives probabilistes : $\mathbf{x} \sim (\vec{\bar{x}}, \Sigma_{\mathbf{xx}})$

- Action d'une transformation : $\mathbf{f} \star \mathbf{x}$,

Implantation (suite)

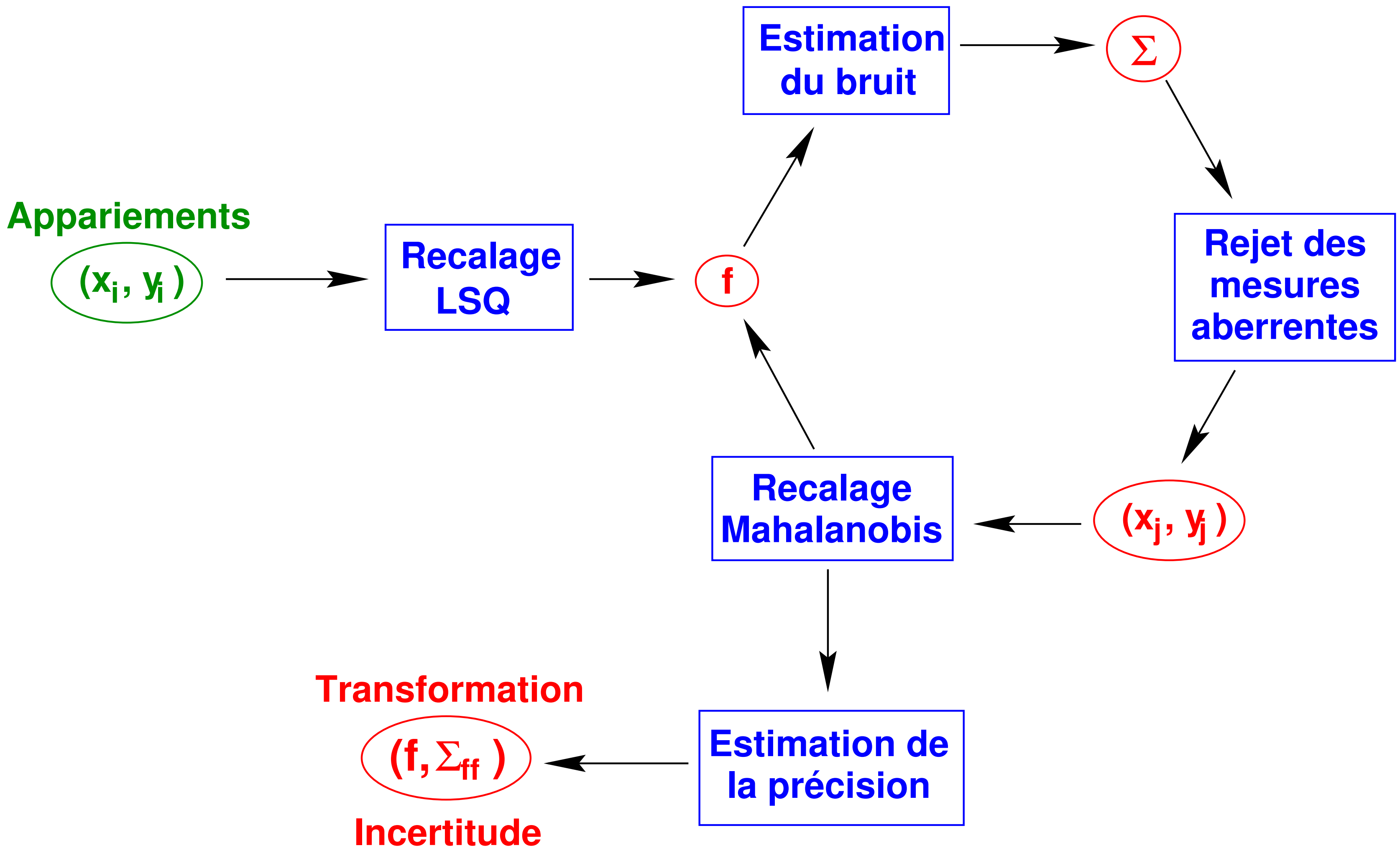
Briques de base

- Distance : $\text{dist}(\mathbf{x}, \mathbf{y}) = \|\vec{f}^{(-1)}_{\vec{x}} \star \vec{y}\|$,
- Mise en place et lecture de la covariance,
- Information et variance,
- Génération aléatoire Gaussienne et uniforme,
- Moyenne et de la covariance d'une population,
- ...

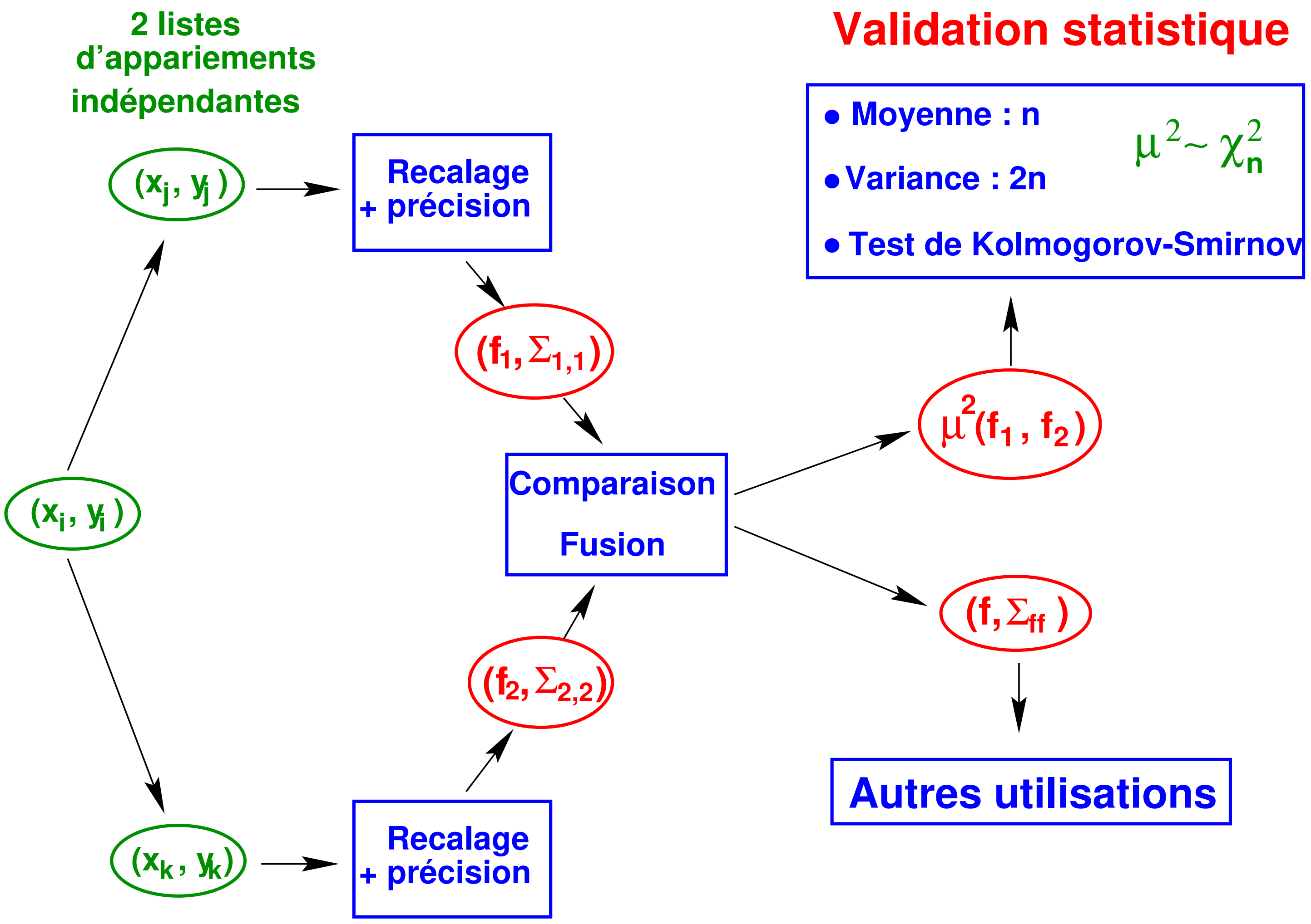
Recalage et fusion

- Critères:
  - Moindres carrés :
  - Distance de Mahalanobis :
- Méthodes
  - Descente de gradient
  - Matrice hessienne : incertitude au minimum.
  - Filtrage de Kalman

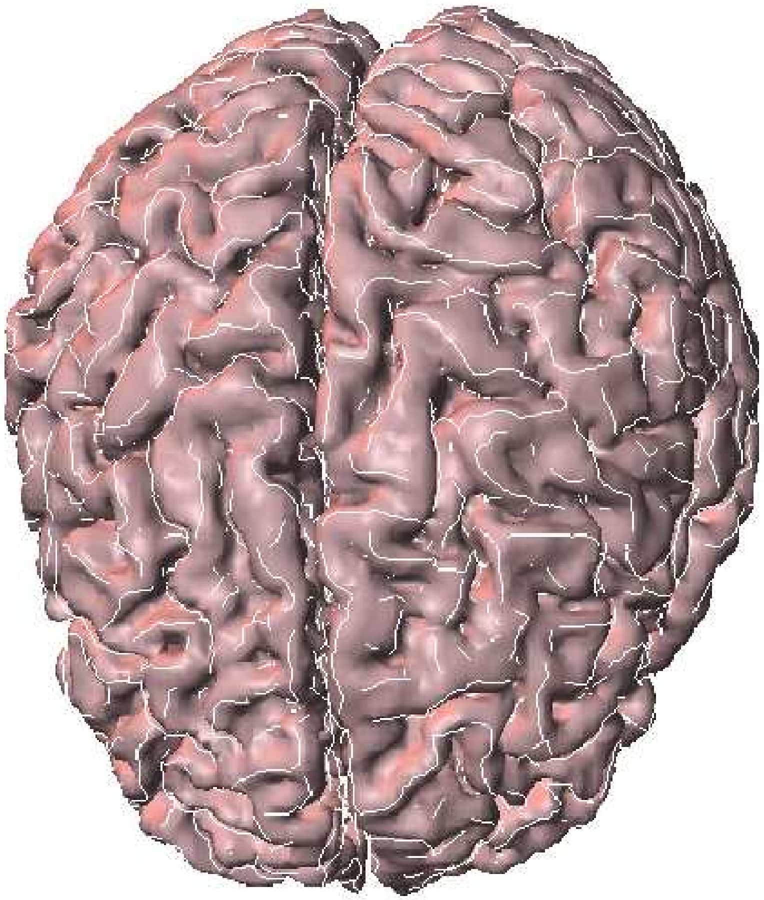
Recalage : algorithme modulaire

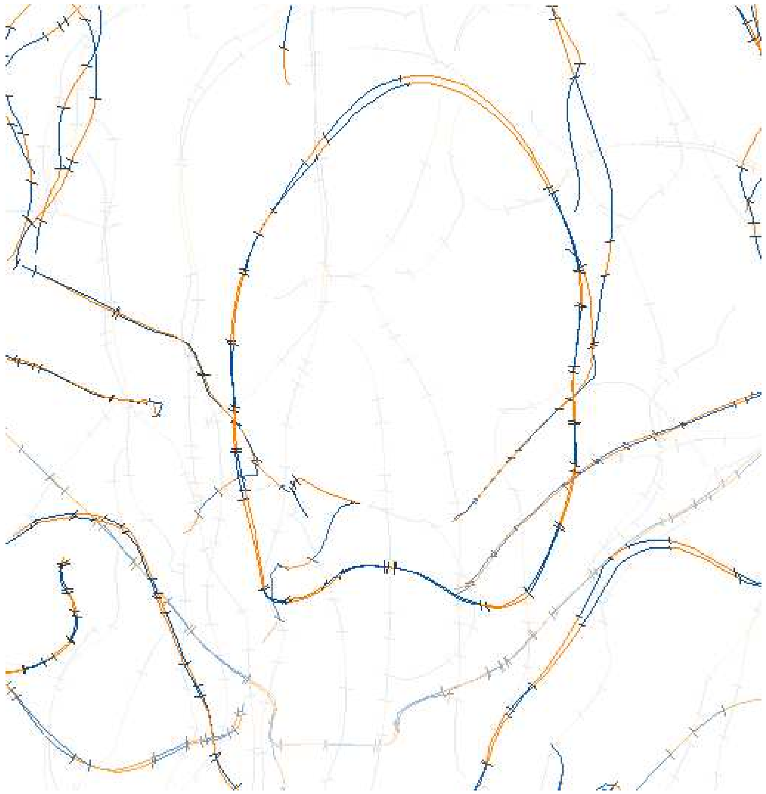


Validation de l'incertitude sur le recalage



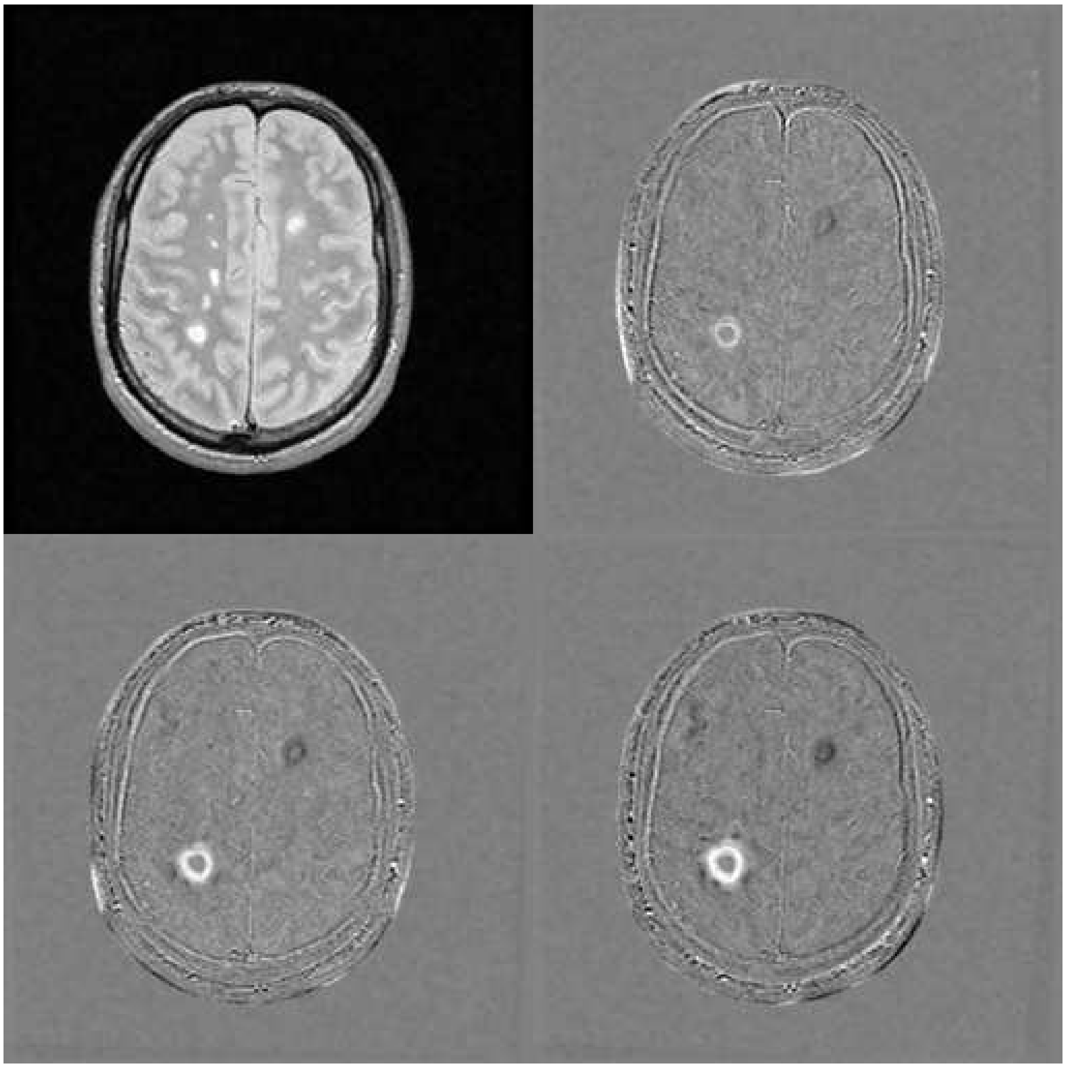
Imagerie médicale : extraction des primitives





Précision du recalage

Évolution de la sclérose en plaque

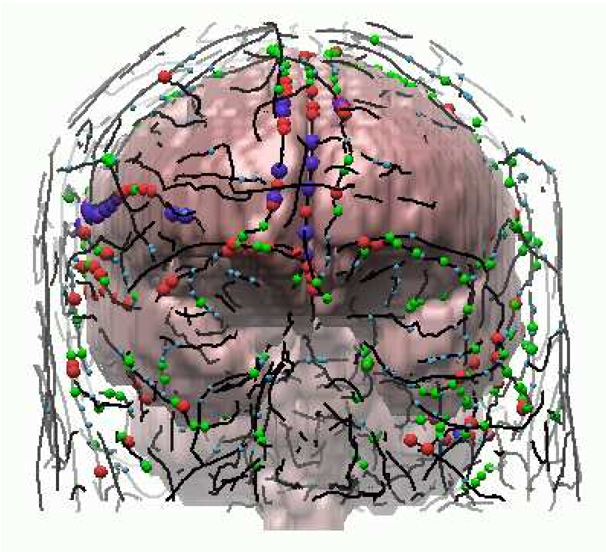


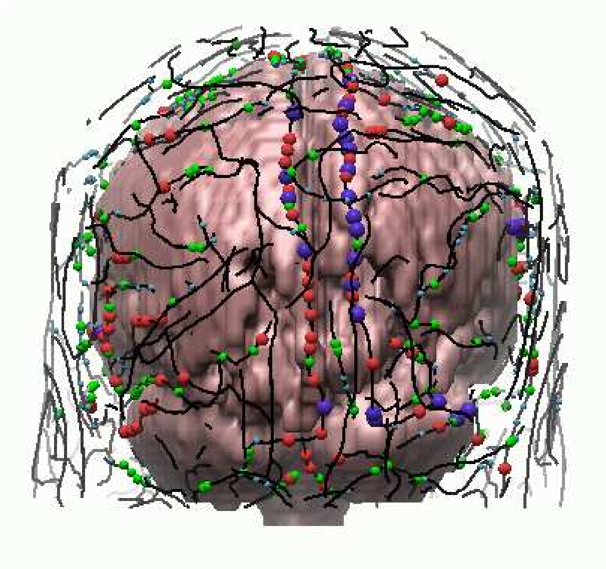
Courtesy of Dr. R. Kikinis

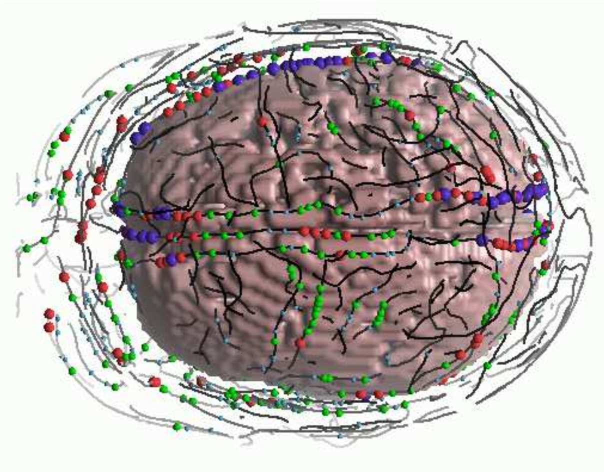
Précision : (Voxel : 1 mm x 1 mm x 3 mm)

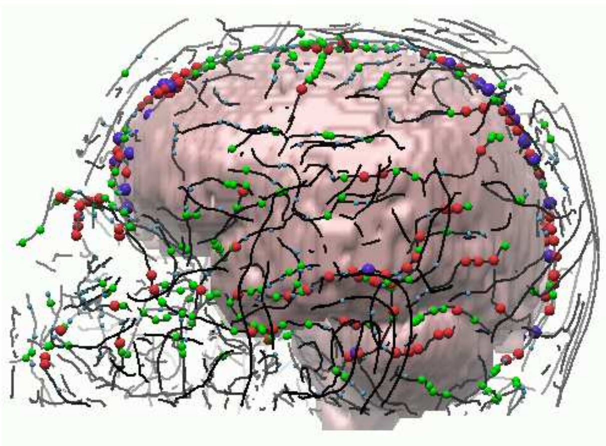
- Écart type aux coins de l'image : 0.114 mm
- Écart type moyen sur les points extrémaux : 0.06 mm

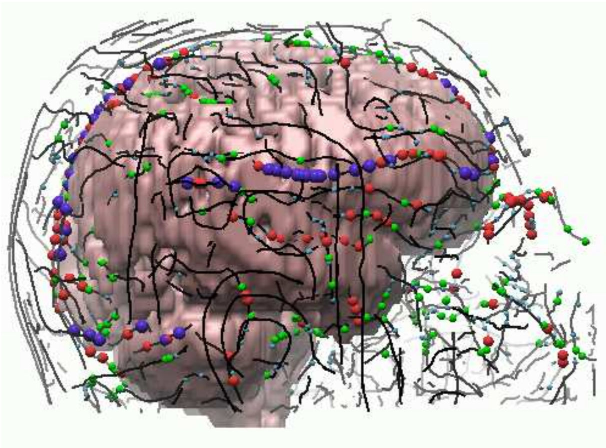
Recalage multiple et modélisation



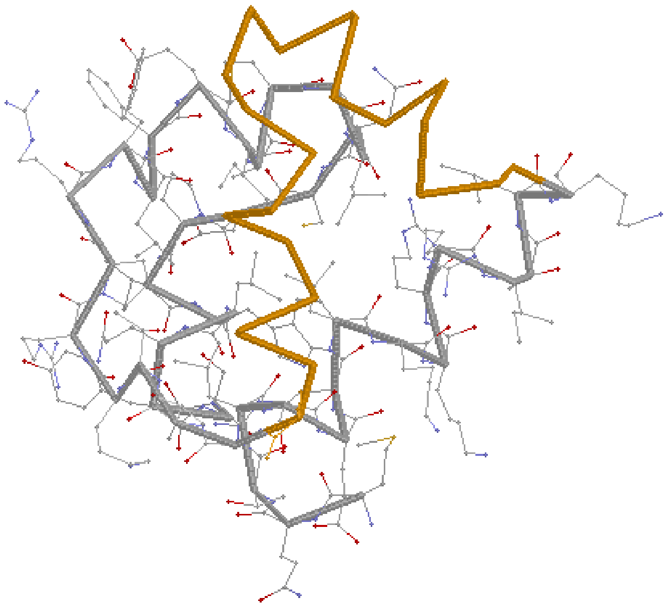




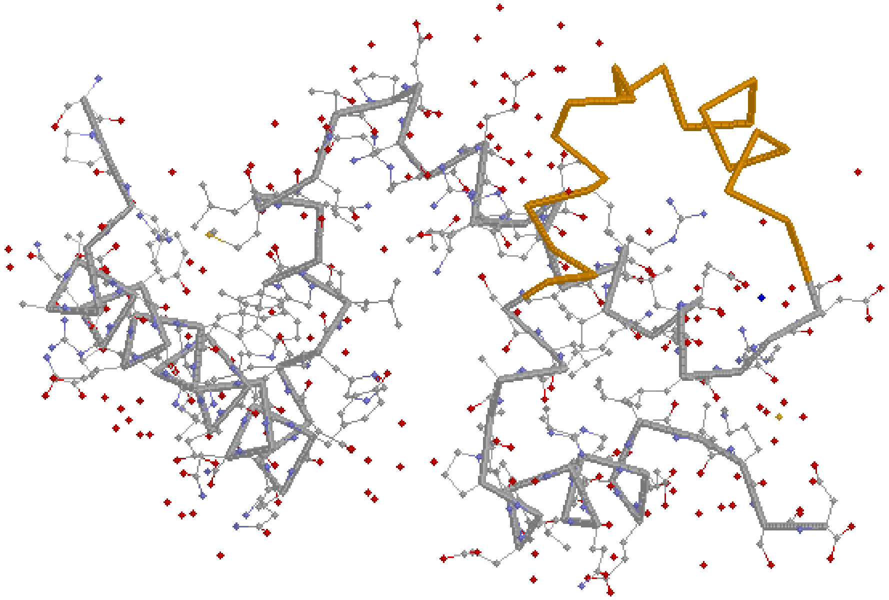




Détection du motif « Helix-turn-Helix »

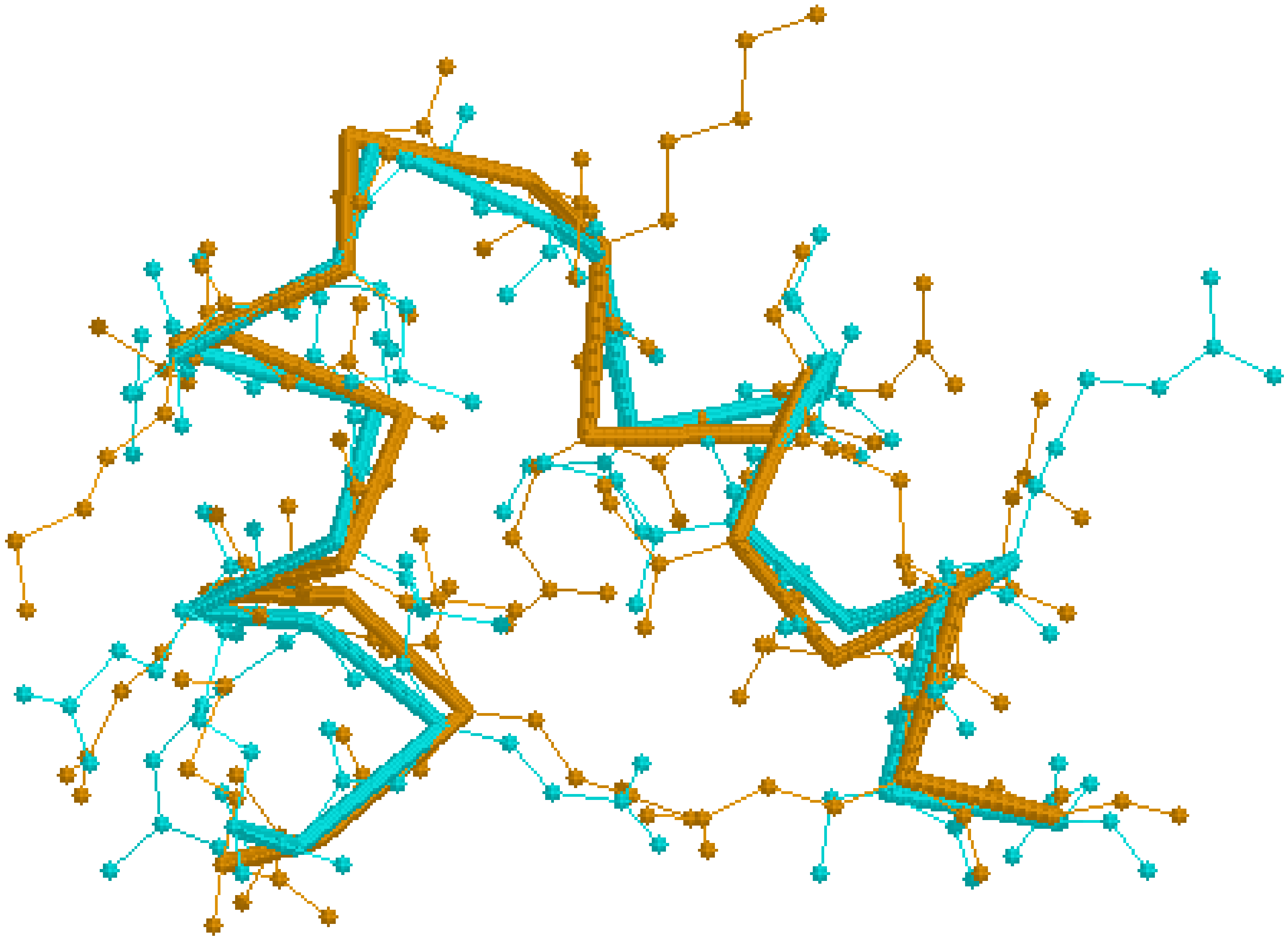


Protéine CRO du phage 434 (2CRO)

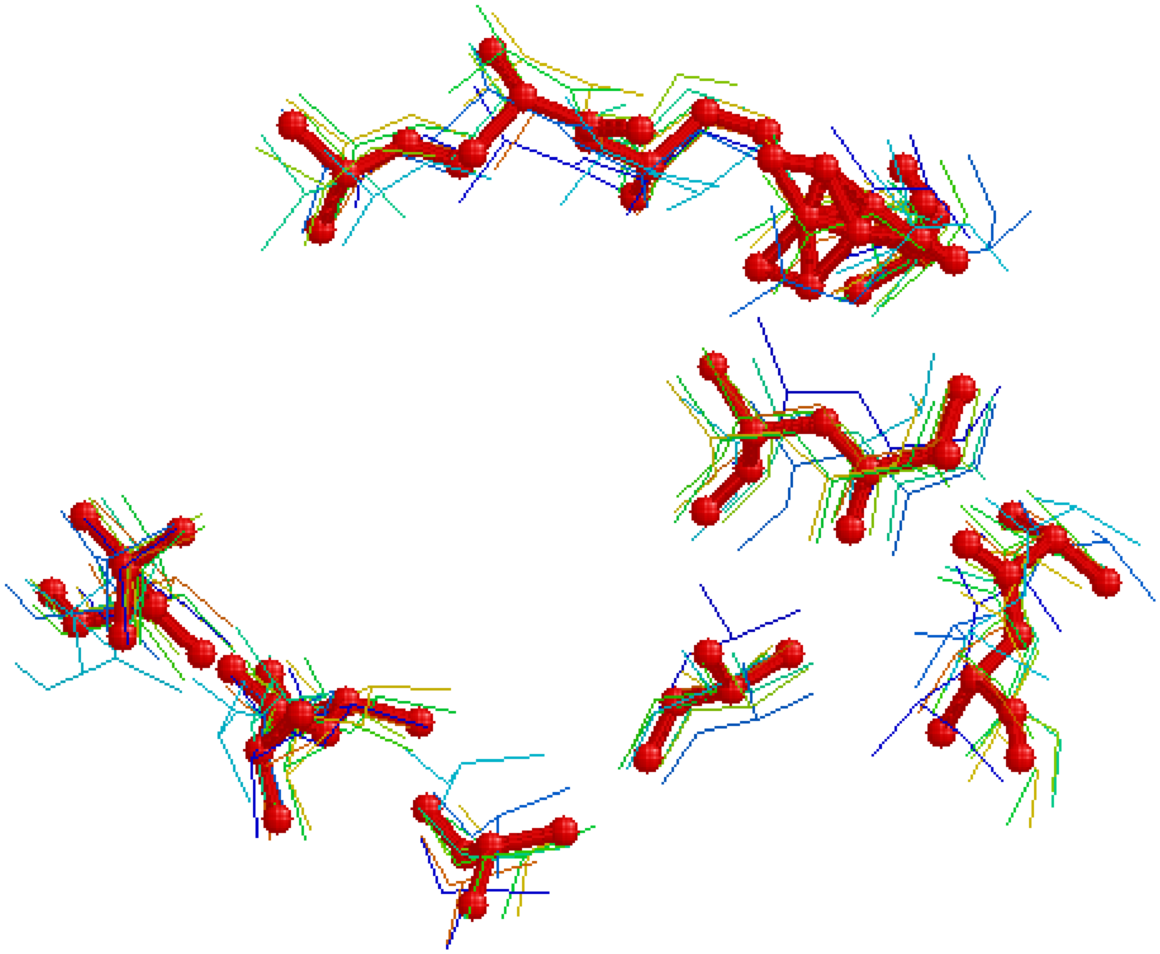


Répresseur Tryptophan de E. Coli (2WRP)

Motif « Helix-turn-Helix »



Modèle du motif de liaison du hème (9 globines)



Conclusion

- Approche théorique rigoureuse de l'incertitude basée sur la géométrie
- Conception de briques algorithmiques génériques
- Implantation efficace et robuste

Perspectives

- Primitive la plus proche (ICP et vérification)
- Construction des invariants n-aires (reconnaissance)
- Extension à des groupes de types affine