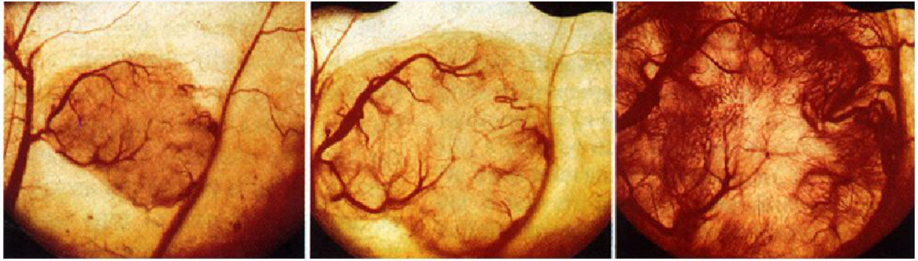
Unsupervised clustering under spatial constraints using multiple equivalence tests

An application to DCE imaging

Fuchen Liu - Université Paris Descartes & Intrasense
Yves Rozenholc - Université Paris Descartes & INRIA Select
Charles-André Cuénod - Université Paris Descartes & HEGP

Workshop
« Traitement, Analyse, Indexation en Imagerie du Vivant »
23-25 juin, Paris

Micro-vascularization

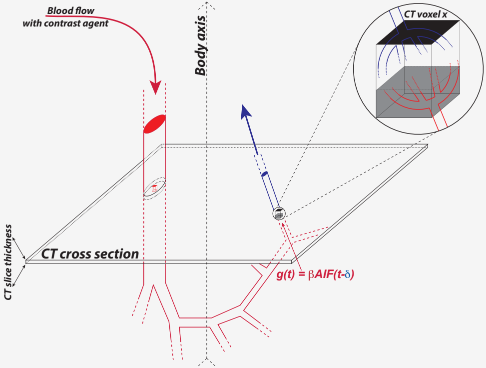


Tumors may have pathologic angiogenesis leading to abnormal vascularization.

Perfusion imaging gives access to functional modification of micro-vascularization.

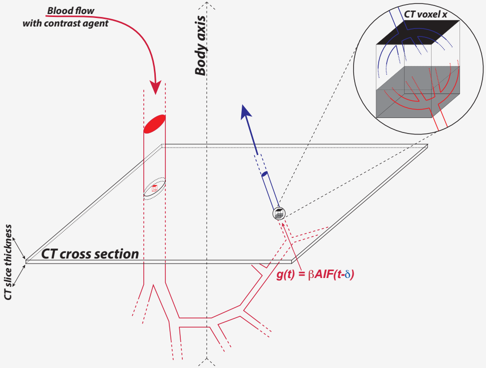
Include DCE-CT, DCE-MRI, or DCE-US.

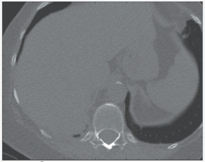
Perfusion imaging
Experimental design



Perfusion imaging

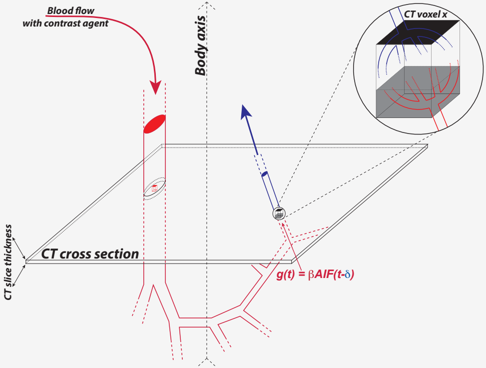
Experimental design

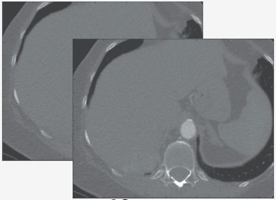




t = 0

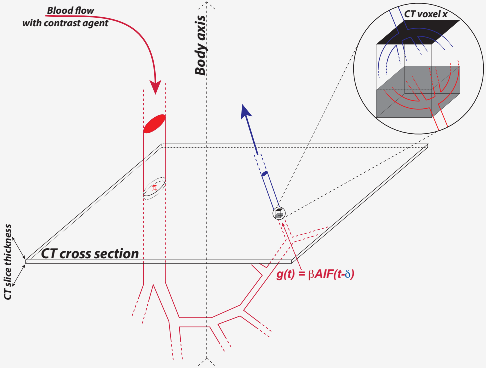
Perfusion imaging
Experimental design

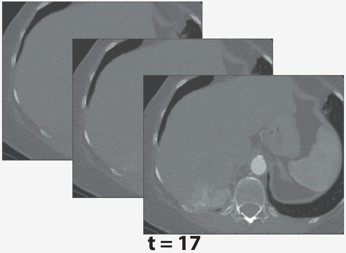




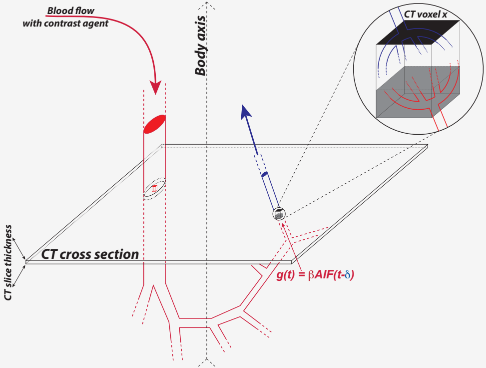
Perfusion imaging

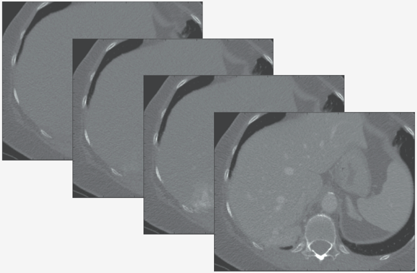
Experimental design



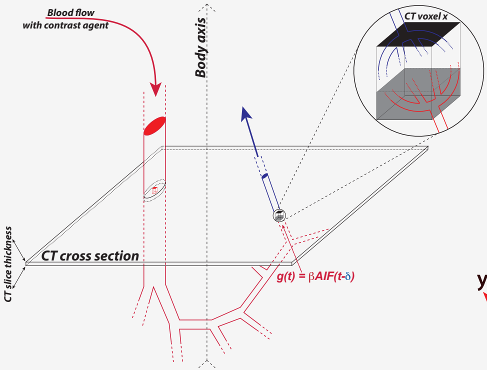


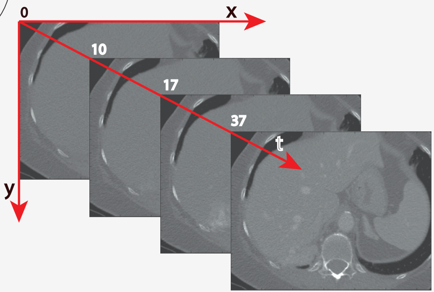
Perfusion imaging
Experimental design



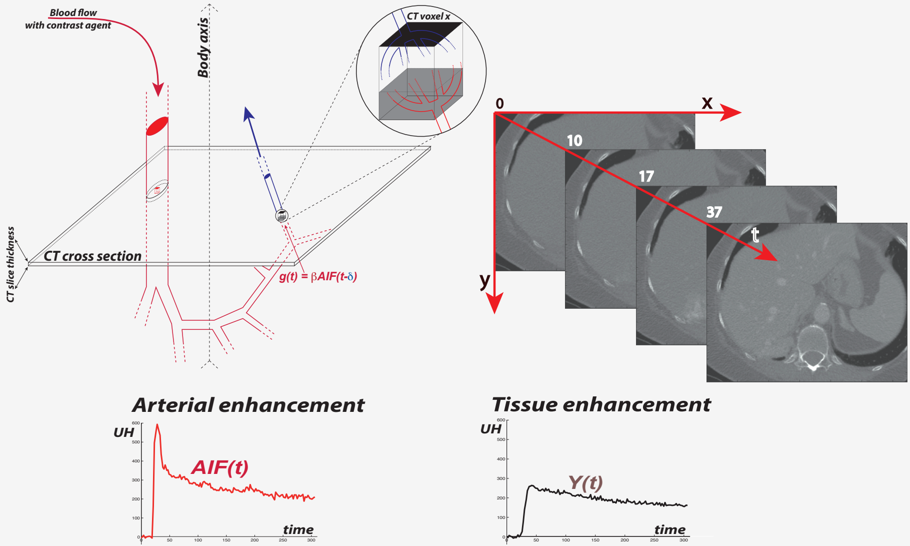


Perfusion imaging
Experimental design





Perfusion imaging
Experimental design



Perfusion analysis

Quantitative analysis uses either

- descriptors from direct evaluation on the dynamic curves;
- compartmental models / system of PDE's;
- blood flow model / Volterra equation of the first type or Laplace deconvolution.

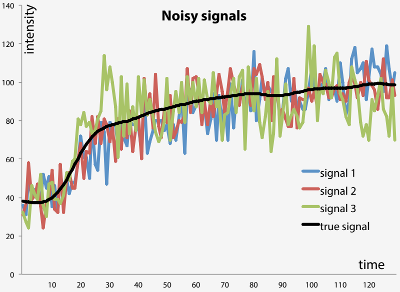
Perfusion analysis

Quantitative analysis uses either

- descriptors from direct evaluation on the dynamic curves;
- compartmental models / system of PDE's;
- blood flow model / Volterra equation of the first type or Laplace deconvolution.

However

all methods suffer from
LOW Signal to Noise Ratio
and
movements

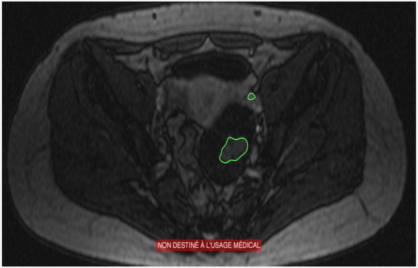


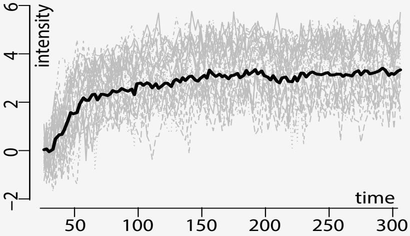
LOW Signal to Noise Ratio is due to instrumental noise of imaging devices:

- X-radiation control in DCE-CT;
- trade-off between spatial and time resolution in DCE-MRI.

Region Of Interest (ROI)

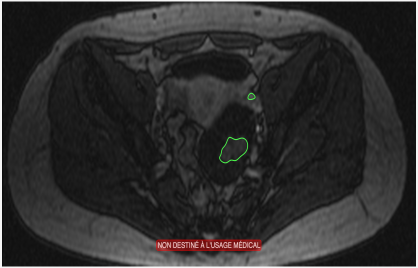
Averaging signals by building ROIs of multiple voxels improves Signal to Noise Ratio

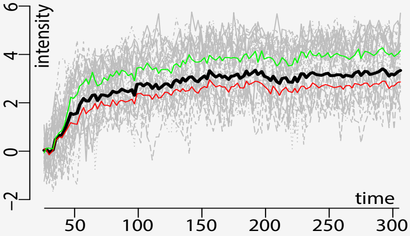




Region Of Interest (ROI)

Averaging signals by building ROIs of multiple voxels improves Signal to Noise Ratio





However may suffer from bad homogeneity

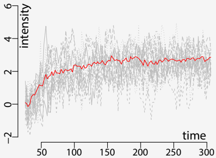
due to mixing different tissues / signals because of :

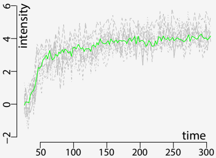
- hard visual perception when drawing manual ROIs ;
- bad comparisons when building automatic ROIs:
  - BAD representation (PCA, basis decomposition, etc) or modelization;
  - BAD choice of distance.

Medical objective

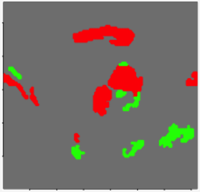
Improve Signal to Noise Ratio by building automatic ROIs

- **controlling temporal homogeneity without prior**





- **controlling spatial homogeneity** $\begin{cases} \textbf{locally} \\ \textbf{at large distance.} \end{cases}$



GOAL: Realize a spatial clusterization which protects temporal structures taking into account that images show neighborhood properties: smoothness, piecewise constant, etc.

Statistical context and statistical objective

Statistical context

We observe a random vector of $\mathbb{R}^n$, $I^x = i^x + \varepsilon^x$ at each location $x \in \mathcal{X}$. Assuming

- ε^x i.i.d. and $\mathbb{E}(\varepsilon^x) = 0$;
- $\mathcal{X} = C_1 \cup \ldots \cup C_K$;
- if $x \in C_k$ for $k = 1, \ldots, K$, then $\mathbb{E}(I^x) := i^x = i_k$;
- if $k \neq \ell$, then $C_k \cap C_\ell = \varnothing$ and $i_k \neq i_\ell$.

Statistical context and statistical objective

Statistical context

We observe a random vector of $\mathbb{R}^n$, $I^x = i^x + \varepsilon^x$ at each location $x \in \mathcal{X}$. Assuming

- ε^x i.i.d. and $\mathbb{E}(\varepsilon^x) = 0$;
- $\mathcal{X} = C_1 \cup \ldots \cup C_K$;
- if $x \in C_k$ for $k = 1, \ldots, K$, then $\mathbb{E}(I^x) := i^x = i_k$;
- if $k \neq \ell$, then $C_k \cap C_\ell = \varnothing$ and $i_k \neq i_\ell$.

Examples:

- $\mathcal{X}$ is a spatial grid of pixels (image) and I^x is a grey-level intensity vector;
- $\mathcal{X}$ is a geographic grid of area and I^x is utility (electricity, water) consumption;
- $\mathcal{X}$ is internet graph and I^x is traffic intensity on node.

Statistical context and statistical objective

Statistical context

We observe a random vector of $\mathbb{R}^n$, $I^x = i^x + \varepsilon^x$ at each location $x \in \mathcal{X}$. Assuming

- ε^x i.i.d. and $\mathbb{E}(\varepsilon^x) = 0$;
- $\mathcal{X} = C_1 \cup \ldots \cup C_K$;
- if $x \in C_k$ for $k = 1, \ldots, K$, then $\mathbb{E}(I^x) := i^x = i_k$;
- if $k \neq \ell$, then $C_k \cap C_\ell = \emptyset$ and $i_k \neq i_\ell$.

Examples:

- $\mathcal{X}$ is a spatial grid of pixels (image) and I^x is a grey-level intensity vector;
- $\mathcal{X}$ is a geographic grid of area and I^x is utility (electricity, water) consumption;
- $\mathcal{X}$ is internet graph and I^x is traffic intensity on node.

Objective

Realize a unsupervised classification to recover the partition $C_1, \ldots, C_K$ such that

$$\boxed{x, y \in \mathcal{X} \text{ belong to the same cluster if and only if } i^x = i^y.}$$

State of the arts of unsupervised classification methods

Main categories:

- **Model-based methods**:
  - Mixture of Gaussians or more complex distributions.
- **Distance-based methods**:
  - K-means;
  - Hierarchical clustering.

Choice of the number K of clusters:

- **penalization techniques like GAP** (Tibshirani, Walther, Hastie, '01)

Actual limitations:

- **Difficulty to select the number K of clusters**:
  - Repeat the same method for different values of K: time consuming;
  - Need assumptions on the distribution inside a cluster (e.g. log-concave).
- **Choice of distance** (e.g. Euclidean distance, Correlation coefficient);
- **Choice of representation** (e.g. PCA, spectral embedding, basis decomposition);
- **Seeds in K-means;**
- **Generic methods which don't take into account spatial regularity.**

Unsupervised classification under spatial constraint: Motivations

Idea 1: Highlight equality

- Build a test such that "=" should be the research hypothesis that is $\mathcal{H}_1$.

Idea 2: Build a universal dissimilarity measure

- Dissimilarity measure derived from the "=" multi-comparison p-value.

Idea 3: Use local spatial homogeneity

- Use neighboring information through a greedy aggregation (low complexity) to take into account the regularity of $x \mapsto i^x$ over the grid $\mathcal{X}$.

Idea 4: Deal with global spatial homogeneity

- Ensure global structure recognition by changing the neighboring information (higher complexity but with less information).

Main tool: Equivalence testing

Notations

Given two subsets X and Y of $\mathcal{X}$, $X \cap Y = \varnothing$,

- Empirical mean over X (resp. Y) is denoted I^X (resp. I^Y);
- $D^{XY} := I^X - I^Y$;
- $d^{XY} := \mathbb{E}(D^{XY})$.

We want to compare i^x and i^y and we test $d^{xy} := i^x - i^y = \vec{0}$ or not.

Main tool: Equivalence testing

Notations

Given two subsets X and Y of $\mathcal{X}$, $X \cap Y = \varnothing$,

- Empirical mean over X (resp. Y) is denoted I^X (resp. I^Y);
- $D^{XY} := I^X - I^Y$;
- $d^{XY} := \mathbb{E}(D^{XY})$.

We want to compare i^x and i^y and we test $d^{xy} := i^x - i^y = \vec{0}$ or not.

Traditional test:

$$\mathcal{H}_0 : d^{xy} = \vec{0} \qquad \text{v.s.} \qquad \mathcal{H}_1 : d^{xy} \neq \vec{0}$$

Non-equivalence is the research hypothesis: **NOT OUR PURPOSE**.

Main tool: Equivalence testing

Notations

Given two subsets X and Y of $\mathcal{X}$, $X \cap Y = \varnothing$,

- Empirical mean over X (resp. Y) is denoted I^X (resp. I^Y);
- $D^{XY} := I^X - I^Y$;
- $d^{XY} := \mathbb{E}(D^{XY})$.

We want to compare i^x and i^y and we test $d^{xy} := i^x - i^y = \vec{0}$ or not.

Traditional test:

$$\mathcal{H}_0 : d^{xy} = \vec{0} \qquad \text{v.s.} \qquad \mathcal{H}_1 : d^{xy} \neq \vec{0}$$

Non-equivalence is the research hypothesis: **NOT OUR PURPOSE**.

Equivalence test: reverse the hypothesis

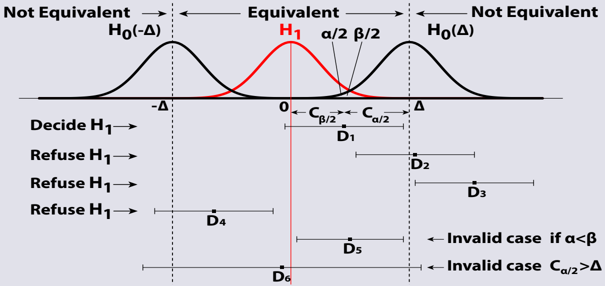
$$\mathcal{H}_0 : d^{xy} \neq \vec{0} \qquad \text{v.s.} \qquad \mathcal{H}_1 : d^{xy} = \vec{0}$$

Equivalence is the research hypothesis: **OUR PURPOSE**.

Equivalence test does not provide strict equivalence but an equivalence up to a given "margin" Δ.

An example of equivalence test: TOST

Gaussian example: Two One-Side Tests (Walker and Nowacki 2011), $n = 1$



Observations not having their confidence interval in $[-\Delta, +\Delta]$ are declared under $\mathcal{H}_1$

Multiple equivalence tests

Multiple equivalence test

Given V_J, $J = 1, \ldots, J_0$ strict linear subspaces of $\mathbb{R}^n$, we consider the **equivalence tests**

$$\mathcal{H}_0^J : d^{XY} \in V_J \setminus \{\vec{0}\} \qquad \text{v.s.} \qquad \mathcal{H}_1 : d^{XY} = \vec{0}.$$

We denote $p_1(X:Y), \ldots, p_{J_0}(X:Y)$ the corresponding **p-values**.

Multiple equivalence tests

Multiple equivalence test

Given V_J, $J = 1, \ldots, J_0$ strict linear subspaces of $\mathbb{R}^n$, we consider the **equivalence tests**

$$\mathcal{H}_0^J : d^{XY} \in V_J \setminus \{\vec{0}\} \qquad \text{v.s.} \qquad \mathcal{H}_1 : d^{XY} = \vec{0}.$$

We denote $p_1(X:Y), \ldots, p_{J_0}(X:Y)$ the corresponding **p-values**.

Intersection-Union Test (IUT)

$$\mathcal{H}_0 = \bigcup_J \mathcal{H}_0^J \qquad \text{v.s.} \qquad \mathcal{H}_1 = \bigcap_J \mathcal{H}_1^J$$

Multiple equivalence tests

Multiple equivalence test

Given V_J, $J = 1, \ldots, J_0$ strict linear subspaces of $\mathbb{R}^n$, we consider the **equivalence tests**

$$\mathcal{H}_0^J : d^{XY} \in V_J \setminus \{\vec{0}\} \qquad \text{v.s.} \qquad \mathcal{H}_1 : d^{XY} = \vec{0}.$$

We denote $p_1(X : Y), \ldots, p_{J_0}(X : Y)$ the corresponding **p-values**.

Intersection-Union Test (IUT)

$$\mathcal{H}_0 = \bigcup_J \mathcal{H}_0^J \qquad \text{v.s.} \qquad \mathcal{H}_1 = \bigcap_J \mathcal{H}_1^J$$

Theorem (Berger and Hsu (1996))

If R_J are α-level rejection regions of $\mathcal{H}_0^J$, $R = \bigcap_J R_J$ is a α-level rejection region for IUT.

Multiple equivalence tests

Multiple equivalence test

Given V_J, $J = 1, \ldots, J_0$ strict linear subspaces of $\mathbb{R}^n$, we consider the **equivalence tests**

$$\mathcal{H}_0^J : d^{XY} \in V_J \setminus \{\vec{0}\} \qquad \text{v.s.} \qquad \mathcal{H}_1 : d^{XY} = \vec{0}.$$

We denote $p_1(X:Y), \ldots, p_{J_0}(X:Y)$ the corresponding **p-values**.

Intersection-Union Test (IUT)

$$\mathcal{H}_0 = \bigcup_J \mathcal{H}_0^J \qquad \text{v.s.} \qquad \mathcal{H}_1 = \bigcap_J \mathcal{H}_1^J$$

Theorem (Berger and Hsu (1996))

If R_J are α-level rejection regions of $\mathcal{H}_0^J$, $R = \bigcap_J R_J$ is a α-level rejection region for IUT.

Corollary

Let p (resp. p_J) be the p-value for $\mathcal{H}_0$ (resp. $\mathcal{H}_0^J$): if $R = \bigcap_J R_J$ then $p = \max_J(p_J)$.

Multiple equivalence tests

Multiple equivalence test

Given V_J, $J = 1, \dots, J_0$ strict linear subspaces of $\mathbb{R}^n$, we consider the **equivalence tests**

$$\mathcal{H}_0^J : d^{XY} \in V_J \setminus \{\vec{0}\} \qquad \text{v.s.} \qquad \mathcal{H}_1 : d^{XY} = \vec{0}.$$

We denote $p_1(X:Y), \dots, p_{J_0}(X:Y)$ the corresponding **p-values**.

Intersection-Union Test (IUT)

$$\mathcal{H}_0 = \bigcup_J \mathcal{H}_0^J \qquad \text{v.s.} \qquad \mathcal{H}_1 = \bigcap_J \mathcal{H}_1^J$$

Theorem (Berger and Hsu (1996))

If R_J are α-level rejection regions of $\mathcal{H}_0^J$, $R = \bigcap_J R_J$ is a α-level rejection region for IUT.

Corollary

Let p (resp. p_J) be the p-value for $\mathcal{H}_0$ (resp. $\mathcal{H}_0^J$): if $R = \bigcap_J R_J$ then $p = \max_J(p_J)$.

Universal dissimilarity measure

The universal dissimilarity measure is $\max_J(p_J)$.

Neighboring clustering with universal dissimilarity measure $\max_J(p_J)$.

An image with 25 pixels

5x5 grid image

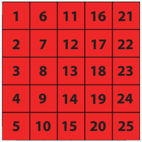
1	6	11	16	21
2	7	12	17	22
3	8	13	18	23
4	9	14	19	24
5	10	15	20	25

- grid $\mathcal{X} := \{1, \ldots, 25\}$.

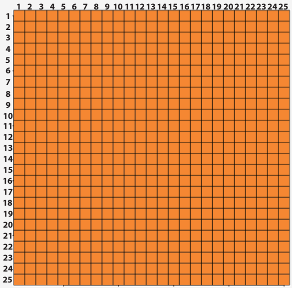
Neighboring clustering with universal dissimilarity measure $\max_J(p_J)$.

An image with 25 pixels

5x5 grid image



25x25 dissimilarity sparse matrix $\mathbf{P}^{(0)}$:

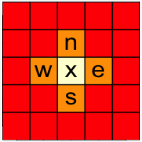


- grid $\mathcal{X} := \{1, \ldots, 25\}$.

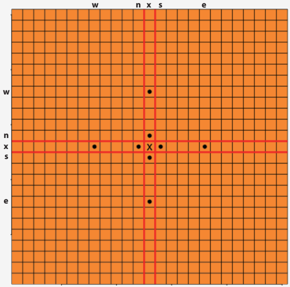
Neighboring clustering with universal dissimilarity measure $\max_J(p_J)$.

Initialization

5x5 grid image



25x25 dissimilarity sparse matrix $\mathbf{P}^{(0)}$:



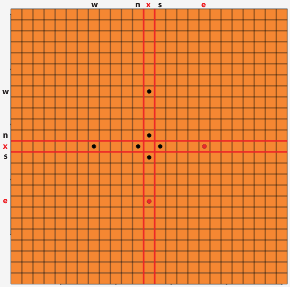
- $\mathcal{P}^{(0)} := \mathcal{X} = \{1, \ldots, 25\}$;
- $\mathcal{V}^{(0)}(x) := \{\{n\}, \{s\}, \{e\}, \{w\}\}$ for $x \in \mathcal{X}$;
- $\mathbf{P}^{(0)}(x, y) := p(x : y)$ for $y \in \mathcal{V}^{(0)}(x)$;
- $\dim(\mathbf{P}^{(0)}) = 25x25$ BUT less than $4x25$ non-zero elements.

Neighboring clustering with universal dissimilarity measure $\max_J(p_J)$.

Iteration t to $t+1$: example of $t = 0$

5x5 grid image

25x25 dissimilarity sparse matrix $\mathbf{P}^{(0)}$:



Optimization step

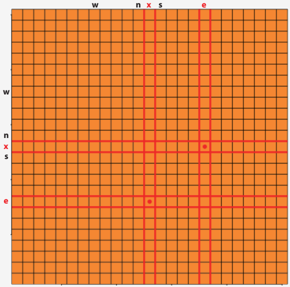
$$(X, Y) = \arg\min(\mathbf{P}^{(t)}) \qquad \text{for } t = 0: \quad X = \{x\} \text{ and } Y = \{e\}$$

Neighboring clustering with universal dissimilarity measure $\max_J(p_J)$.

Iteration t to $t+1$: example of $t = 0$

5x5 grid image

25x25 dissimilarity sparse matrix $\mathbf{P}^{(0)}$:



Optimization step

$$(X, Y) = \arg\min(\mathbf{P}^{(t)}) \qquad \text{for } t = 0: \quad X = \{x\} \text{ and } Y = \{e\}$$

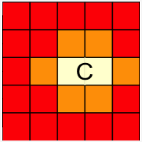
Update phase 1 - downsizing by 2

Remove two columns and two lines of $\mathbf{P}^{(t)}$ corresponding to X and Y.

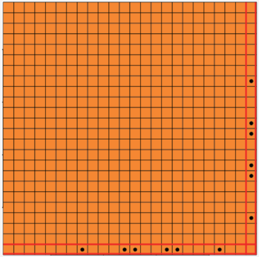
Neighboring clustering with universal dissimilarity measure $\max_J(p_J)$.

Iteration t to $t+1$: example of $t = 0$

5x5 grid image



24x24 dissimilarity sparse matrix $\mathbf{P}^{(1)}$:



Update phase 2 - upsizing by 1: Update partition, neighbors and dissimilarity matrix

- $C := X \cup Y$; **new cluster**
- $\mathcal{P}^{(t+1)} := \mathcal{P}^{(t)} \setminus X \setminus Y \cup C$; **new partition**
- $\mathcal{V}^{(t+1)}(C) := \mathcal{V}^{(t)}(X) \cup \mathcal{V}^{(t)}(Y) \setminus X \setminus Y$; **new neighbors**
- $\boxed{\mathbf{P}^{(t+1)}(C, Z) = p(C : Z) \text{ for } Z \in \mathcal{V}^{(t+1)}(C)}$ **new dissimilarities**

From local to global clustering with universal dissimilarity measure

Local clustering — uses spatial regularity

- Starts from a partition made of all voxels as singletons;
- Builds, iteratively, successive partitions by aggregating two clusters at each step;
- Provides a hierarchical sequence of partitions with decreasing sizes from N to 1.

From local to global clustering with universal dissimilarity measure

Local clustering — uses spatial regularity

- Starts from a partition made of all voxels as singletons;
- Builds, iteratively, successive partitions by aggregating two clusters at each step;
- Provides a hierarchical sequence of partitions with decreasing sizes from N to 1.

Global clustering

When local clustering stops, process continues by changing the neighborhood structure:

- neighbors of neighbors

$$\mathcal{V}(X) := \bigcup_{Y \in \mathcal{V}(X)} \mathcal{V}(Y) \setminus \{X\}.$$

- all other clusters

$$\mathcal{V}(X) := \mathcal{P} \setminus \{X\}.$$

Automatic selection of number of clusters

With s clusters, we fix type I error to α and search for $c_\alpha(s)$ s.t.

$$\mathbb{P}[\min \mathbf{P}^{(N-s)} \leqslant c_\alpha(s)] = \mathbb{P}\left[\min_{1\leqslant k,\ell\leqslant s}\left\{p(C_k^{(N-s)} : C_\ell^{(N-s)})\right\} \leqslant c_\alpha(s)\right] = \alpha$$

Automatic selection of number of clusters

With s clusters, we fix type I error to α and search for $c_\alpha(s)$ s.t.

$$\mathbb{P}[\min \mathbf{P}^{(N-s)} \leqslant c_\alpha(s)] = \mathbb{P}\left[\min_{1\leqslant k,\ell\leqslant s}\left\{p(C_k^{(N-s)} : C_\ell^{(N-s)})\right\} \leqslant c_\alpha(s)\right] = \alpha$$

As $p(X : Y) := \max_{J=1,\dots,J_0} p_J(X : Y)$

Control function

If p-values p_J are uniform (abs. cont. tests), assuming independence of the p-values:

$$c_\alpha(s) = (1-(1-\alpha)^{2/s(s-1)})^{1/J_0}, \quad \text{for all } s = N, \dots, 1.$$

Automatic selection of number of clusters

With s clusters, we fix type I error to α and search for $c_\alpha(s)$ s.t.

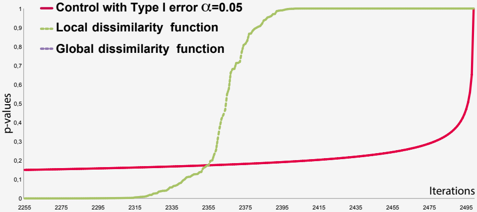
$$\mathbb{P}[\min \mathbf{P}^{(N-s)} \leqslant c_\alpha(s)] = \mathbb{P}\left[\min_{1\leqslant k,\ell\leqslant s}\left\{p(C_k^{(N-s)} : C_\ell^{(N-s)})\right\} \leqslant c_\alpha(s)\right] = \alpha$$

$$\text{As } p(X:Y) := \max_{J=1,\dots,J_0} p_J(X:Y)$$

Control function

If p-values p_J are uniform (abs. cont. tests), assuming independence of the p-values:

$$c_\alpha(s) = (1-(1-\alpha)^{2/s(s-1)})^{1/J_0}, \quad \text{for all } s = N, \dots, 1.$$



Note that $c_\alpha(.)$ does not depend on Δ. Only p-values depend on Δ.

Automatic selection of number of clusters

With s clusters, we fix type I error to α and search for $c_\alpha(s)$ s.t.

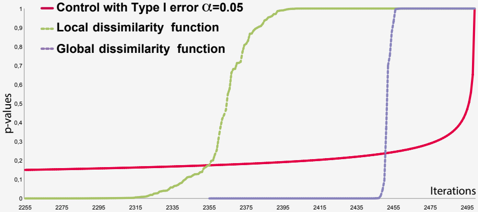
$$\mathbb{P}[\min \mathbf{P}^{(N-s)} \leqslant c_\alpha(s)] = \mathbb{P}\left[\min_{1\leqslant k,\ell\leqslant s}\left\{p(C_k^{(N-s)} : C_\ell^{(N-s)})\right\} \leqslant c_\alpha(s)\right] = \alpha$$

$$\text{As } p(X:Y) := \max_{J=1,\ldots,J_0} p_J(X:Y)$$

Control function

If p-values p_J are uniform (continuous tests), assuming independence of the p-values:

$$c_\alpha(s) = (1-(1-\alpha)^{2/s(s-1)})^{1/J_0}, \quad \text{for all } s = N, \ldots, 1.$$



Note that $c_\alpha(.)$ does not depend on Δ. Only p-values depend on Δ.

Statistical framework

Notations

Given two sets of voxels X and Y, $X \cap Y = \varnothing$

$$D^{XY} := \frac{\frac{1}{|X|}\sum_{x\in X} I^x - \frac{1}{|Y|}\sum_{y\in Y} I^y}{\rho(X:Y)} \sim \mathcal{N}(d^{XY}, Id_n),$$

where

$$d^{XY} := \frac{\frac{1}{|X|}\sum_{x\in X} i^x - \frac{1}{|Y|}\sum_{y\in Y} i^y}{\rho(X:Y)}, \qquad \rho^2(X:Y) := \frac{1}{|X|} + \frac{1}{|Y|}.$$

$\Pi_J(D^{XY})$ **denotes the projection of** D^{XY} **over the** 2^J**-bins piecewise constant vectors.**

Multi-resolution equivalence tests for dynamics comparison

$$\mathcal{H}_0^J : \|\Pi_J(D^{XY})\|^2 \neq 0 \quad \text{v.s.} \quad \mathcal{H}_1 : \|\Pi_J(D^{XY})\|^2 = 0.$$

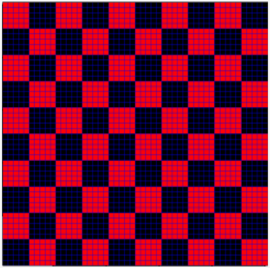
in this case :

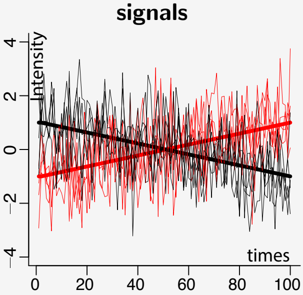
$$p_J(X:Y) := \mathbb{P}\left(\chi^2(2^J, \Delta^2 n) \leqslant \|\Pi_J(D^{XY})\|^2\right).$$

see Baraud et al (2003) for traditional test.

Simulation: chess board

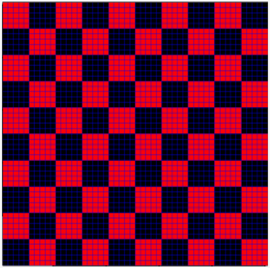
2 clusters in a 50×50 image



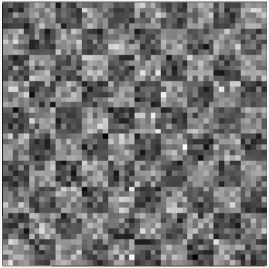


Simulation: chess board

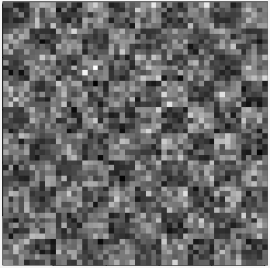
2 clusters in a 50×50 image



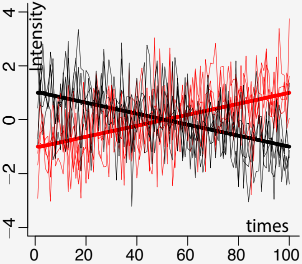
$t = 1$



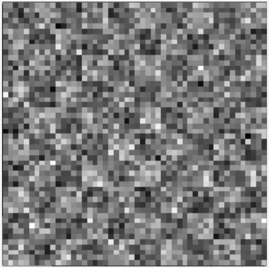
$t = 10$



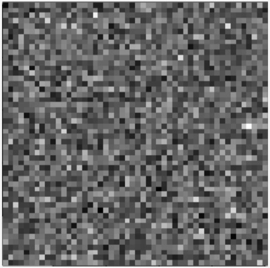
signals



$t = 20$



$t = 50$



Simulation: chess board

Data information

- Spatial resolution: 50×50 voxels;
- Temporal resolution: 100 times points;
- Noise level: $\sigma = 1$.

Parameter setting

- $\Delta = \mathbf{1}$; margin
- $\alpha = 0.05$. Type I error

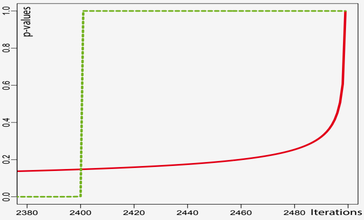
Simulation: chess board

Data information

- Spatial resolution: 50×50 voxels;
- Temporal resolution: 100 times points;
- Noise level: $\sigma = 1$.

Parameter setting

- $\Delta = \mathbf{1}$; margin
- $\alpha = 0.05$. Type I error



Type I error control

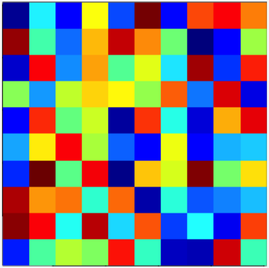
Simulation: chess board

Data information

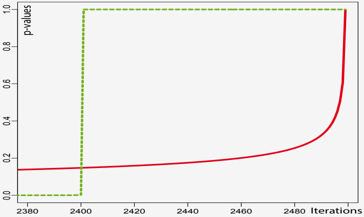
- Spatial resolution: 50×50 voxels;
- Temporal resolution: 100 times points;
- Noise level: $\sigma = 1$.

Parameter setting

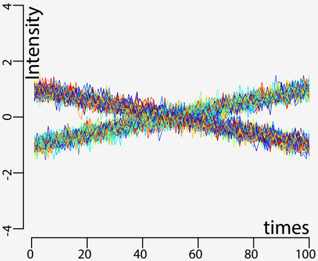
- $\Delta = \mathbf{1}$; margin
- $\alpha = 0.05$. Type I error



Local step: 100 clusters



Type I error control



Estimated curves

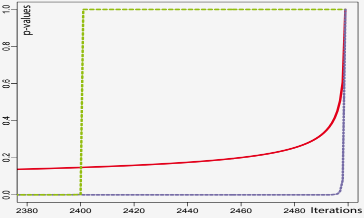
Simulation: chess board

Data information

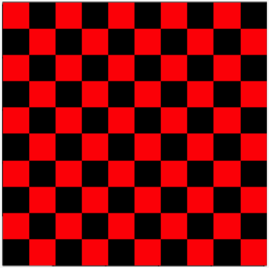
- Spatial resolution: 50×50 voxels;
- Temporal resolution: 100 times points;
- Noise level: $\sigma = 1$.

Parameter setting

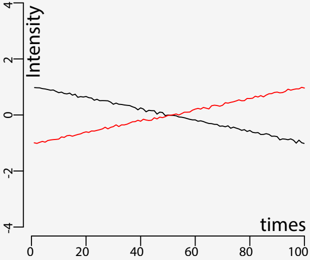
- $\Delta = \mathbf{1}$; margin
- $\alpha = 0.05$. Type I error



Type I error control



Global step: 2 clusters

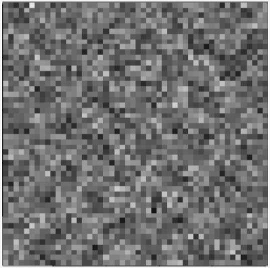


Estimated curves

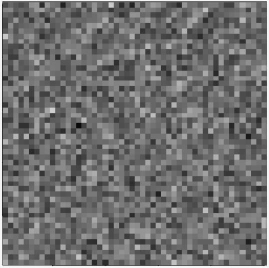
Simulation: chess board

A much harder case

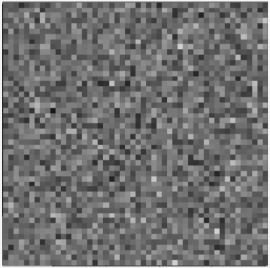
$t = 1$



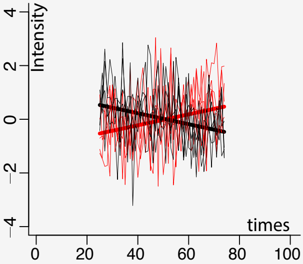
$t = 13$



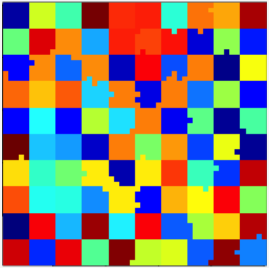
$t = 26$



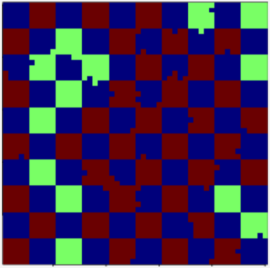
Shrinken signals



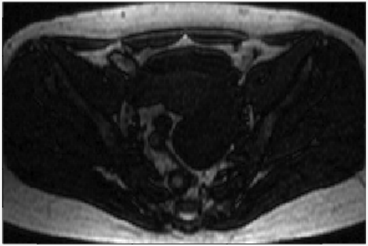
Local: 89 clusters



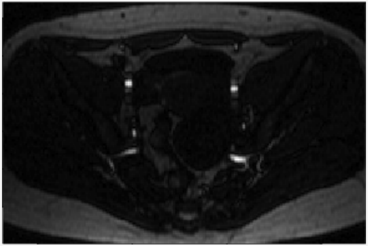
Global: 3 clusters



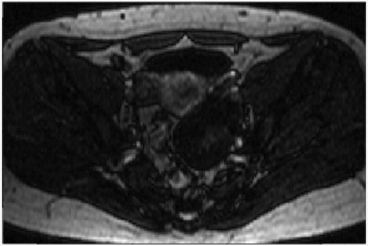
DCE-MRI: Ovarian malignant tumor (192×128 voxels, 119 times points)



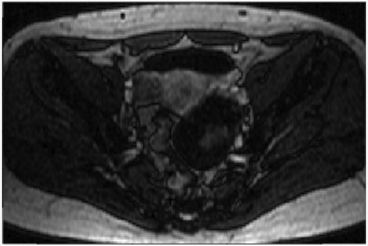
$t = 2$



$t = 14$

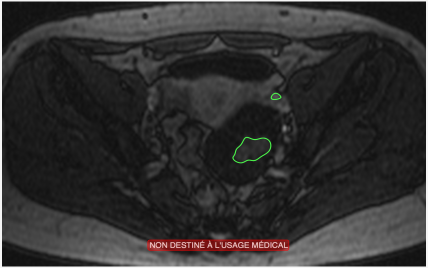


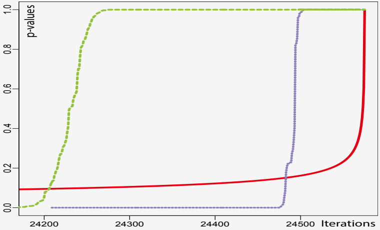
$t = 20$



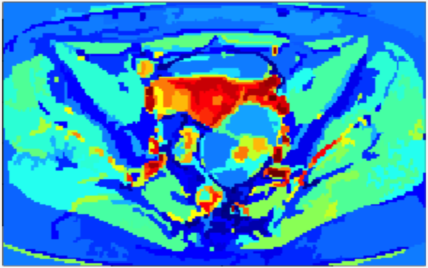
$t = 40$

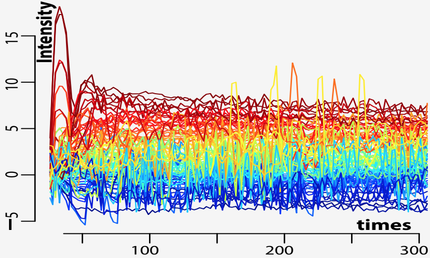
DCE-MRI: Ovarian malignant tumor ($\Delta = \mathbf{3}^2 \times 119$, $\alpha = 0.05$)





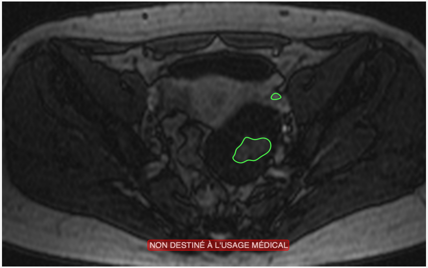
Type I error control

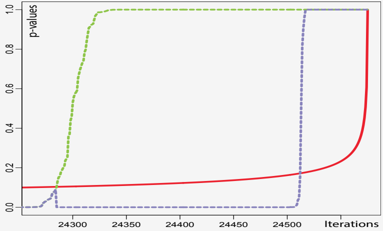




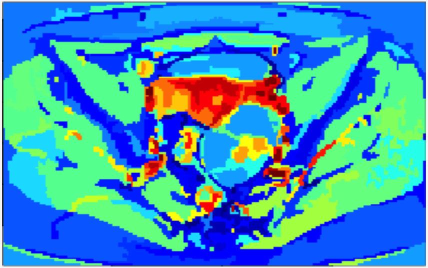
Global step: 94 clusters

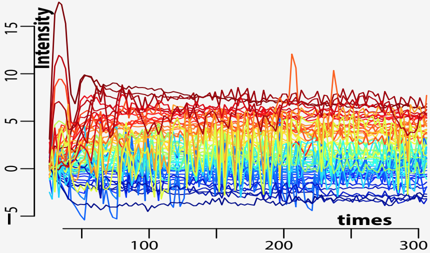
DCE-MRI: Ovarian malignant tumor ($\Delta = \mathbf{3.5}^2 \times 119$, $\alpha = 0.05$)





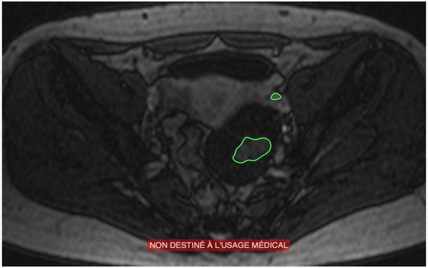
Type I error control

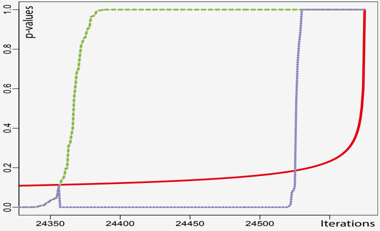




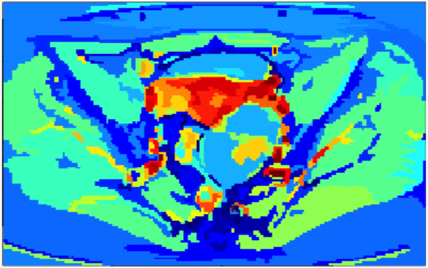
Global step: 64 clusters

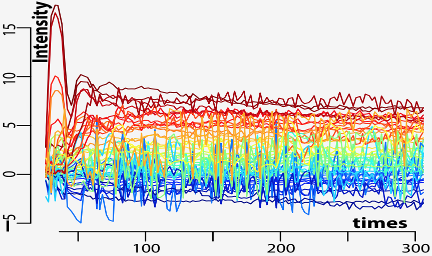
DCE-MRI: Ovarian malignant tumor ($\Delta = \mathbf{4}^2 \times 119$, $\alpha = 0.05$)





Type I error control



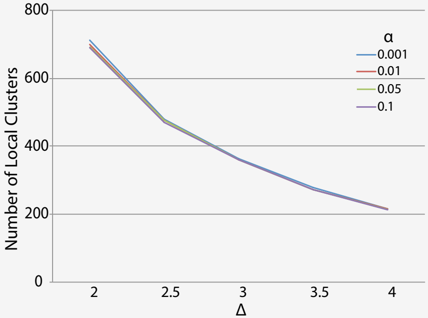


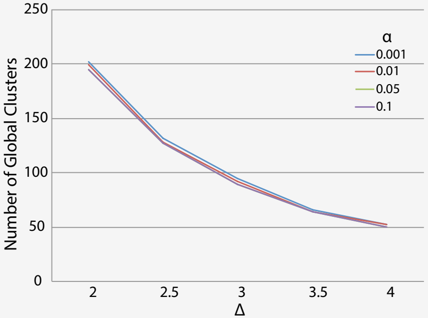
Global step: 51 clusters

Influence of the parameters

Parameter tuning

- Δ : from 2 to 4;
- α : from 0.001 to 0.1;





Only Δ has a major contribution on the number of clusters !

References

- Baraud, Y., Huet, S., and Laurent, B. (2003). Adaptive tests of linear hypotheses by model selection. *Ann. Statist.* **31**, 225–51.
- Berger, R. L. and Hsu, J. C. (1996). Bioequivalence Trials, Intersection-Union Tests and Equivalence Confidence Sets. *Statistical Science.* **11**, 283–319.
- Durot, C. and Rozenholc, Y. (2006). An adaptive test for zero mean. Math. Methods Statist. *Math. Methods Statist.* **15**, 26–60.
- Walker, E. and Nowacki, A. S. (2011). Understanding equivalence and noninferiority testing. *J Gen Intern Med.* **26**, 192–196.
- Rozenholc, Y., Reiß, M., Balvay, D. and Cuénod, C. A. (2010). Growing time-homogeneous neighborhoods for denoising and clustering Dynamic Contrast Enhanced-CT sequences. `https://hal.archives-ouvertes.fr/hal-00712348`.
- Tibshirani, R., Walther, G. and Hastie, T. (2001) Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology).* **63**, 411–423.