Lesion Segmentation using a Spatially Regularized Mixture Model

B. Ozenne[1], F. Subtil[1], L. Østergaard[2], D. Maucort-Boulch[1]

brice.ozenne@chu-lyon.fr

[1]Department of Biostatistics, Hospices Civils de Lyon, France

[2]Center of Functionally Integrative Neuroscience, Århus University,Denmark

Applied statistics - September 23, 2014

Lesion segmentation

Motivations

- **clinical practice :** diagnosis of stroke, multiple sclerosis.
- **clinical research :** objective assessment of the disease.
  - ▷ gold standard for predictive models, drug evaluation.

Limits of manual segmentation

- time consuming.
- source of inter-observer variability.
- difficult in case of complex 3D structures.

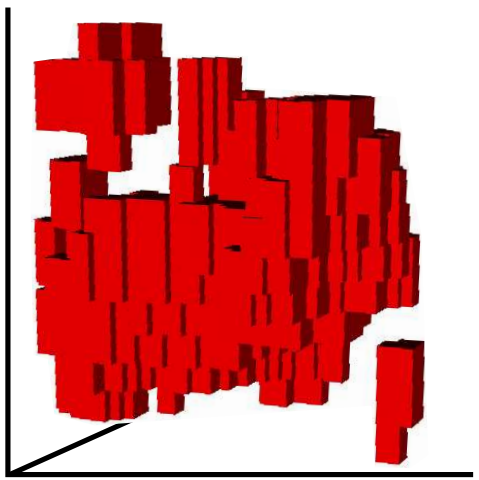


Figure: 3D stroke lesion

State of the art

Current approaches

- level set models

  Osher and Fedkiw, 2003; Weinman et al., 2003; Mouridsen et al., 2013

- supervised learning (glm, machine learning)

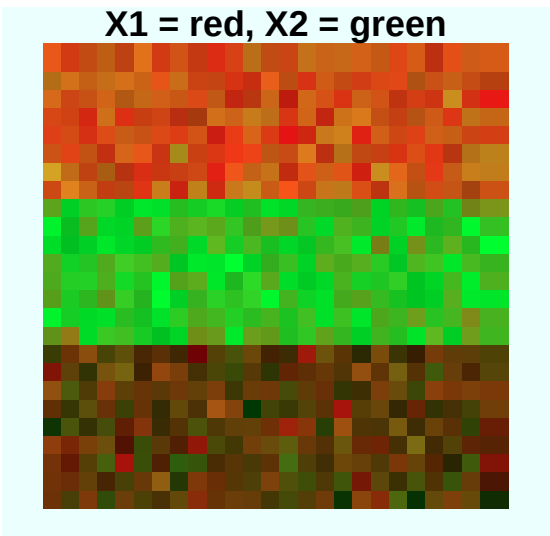
  Klëppel et al., 2011; Sweeney et al., 2013

- finite Mixture Models : very popular
  - ▷ unsupervised
  - ▷ few parameters
  - ▷ flexible modelling framework

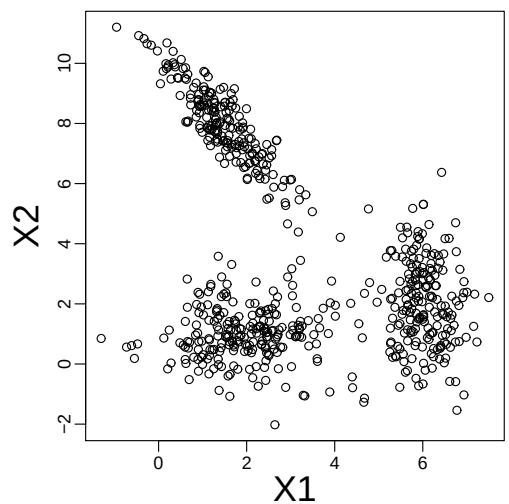
State of the art

Current approaches

- finite Mixture Models : very popular
  - ▷ unsupervised
  - ▷ few parameters
  - ▷ flexible modelling framework



Image



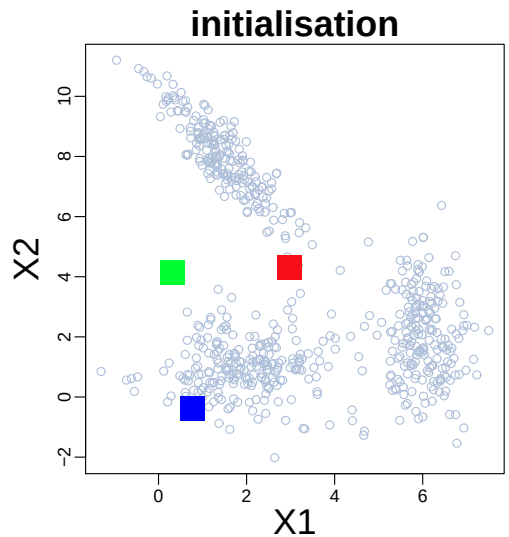
Intensity space

State of the art

Current approaches

- finite Mixture Models : very popular
  - ▷ unsupervised
  - ▷ few parameters
  - ▷ flexible modelling framework

Image

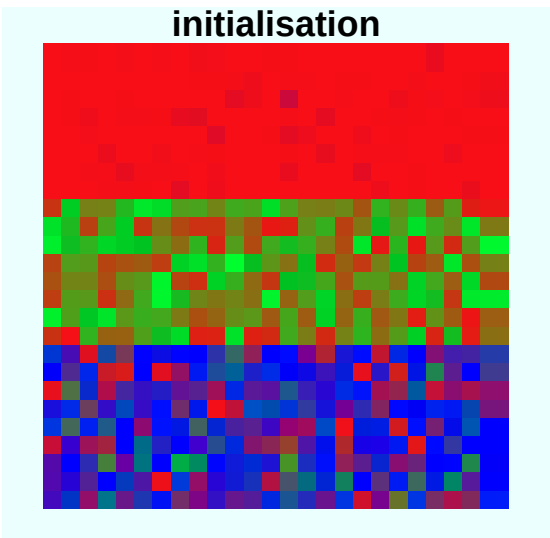


Intensity space

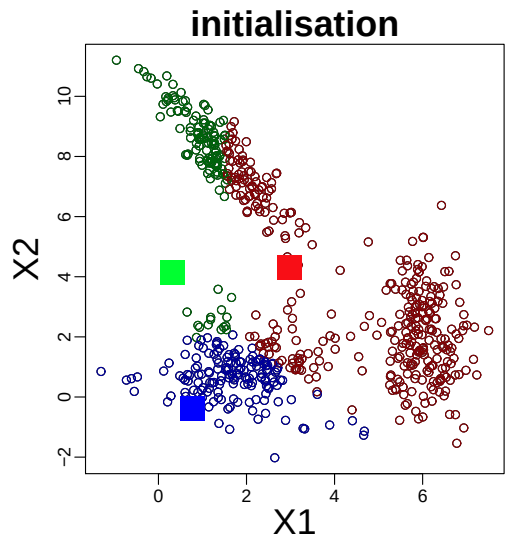
State of the art

Current approaches

- finite Mixture Models : very popular
  - ▷ unsupervised
  - ▷ few parameters
  - ▷ flexible modelling framework



Image

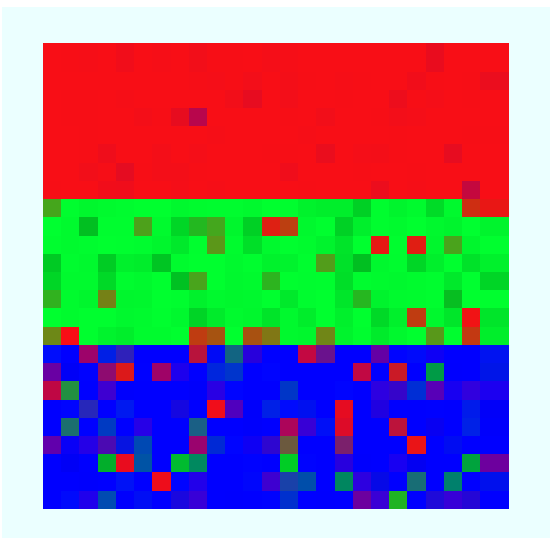


Intensity space

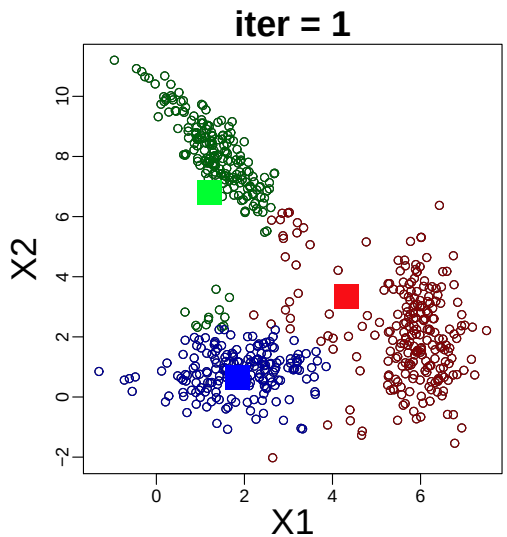
State of the art

Current approaches

- finite Mixture Models : very popular
  - ▷ unsupervised
  - ▷ few parameters
  - ▷ flexible modelling framework



Image

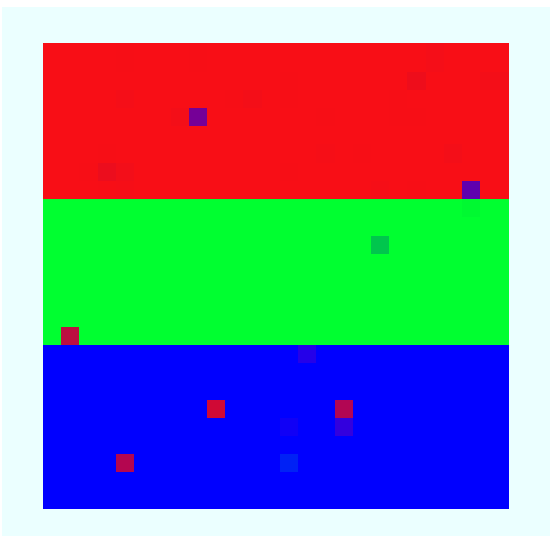


Intensity space

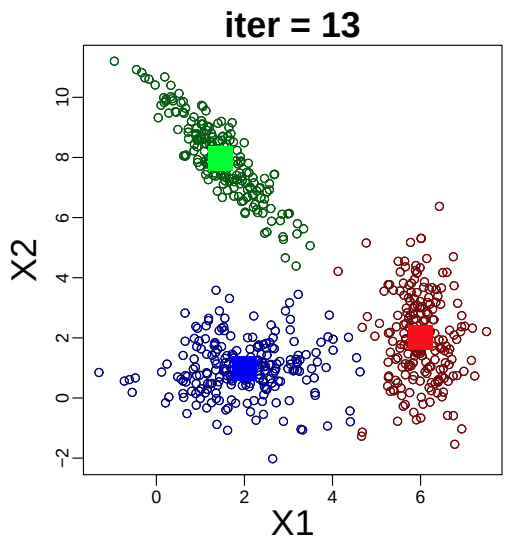
State of the art

Current approaches

- finite Mixture Models : very popular
  - ▷ unsupervised
  - ▷ few parameters
  - ▷ flexible modelling framework



Image



Intensity space

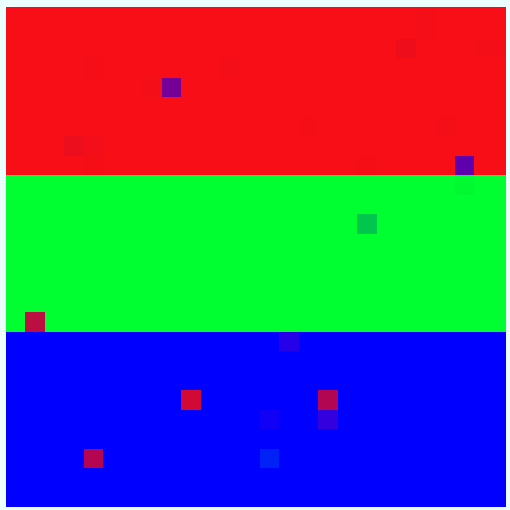
State of the art

Limits

- noise degrades the segmentation
  ▷ *univariate* spatial fMM
  (Woolrich et al., 2005 ; Feng, Tierney, and Magnotta, 2012 ; Zhang et al., 2008)

- white matter disease can be confused with stroke lesion
  ▷ lead to segmentation errors
  ▷ volume over-estimation

⇒ **need for a regional approach**



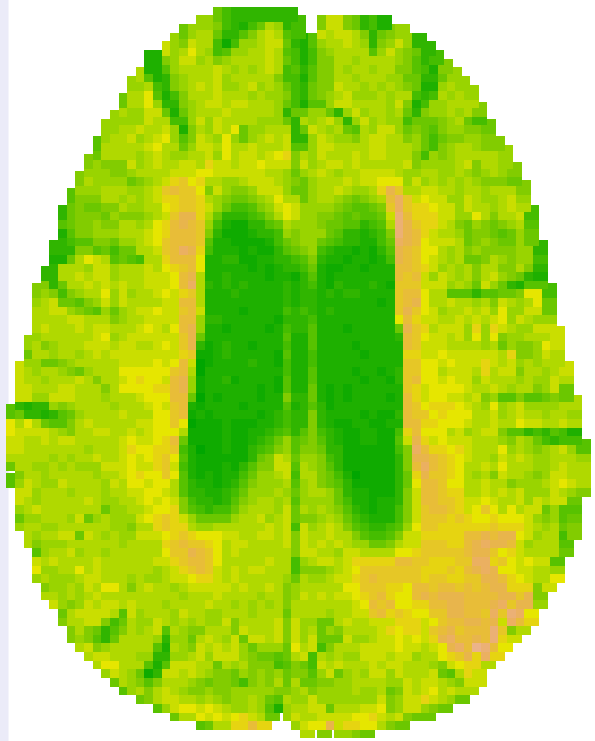
State of the art

Limits

- noise degrades the segmentation
  ▷ *univariate* spatial fMM
  (Woolrich et al., 2005 ; Feng, Tierney, and Magnotta, 2012 ; Zhang et al., 2008)

- white matter disease can be confused with stroke lesion
  ▷ lead to segmentation errors
  ▷ volume over-estimation

⇒ **need for a regional approach**



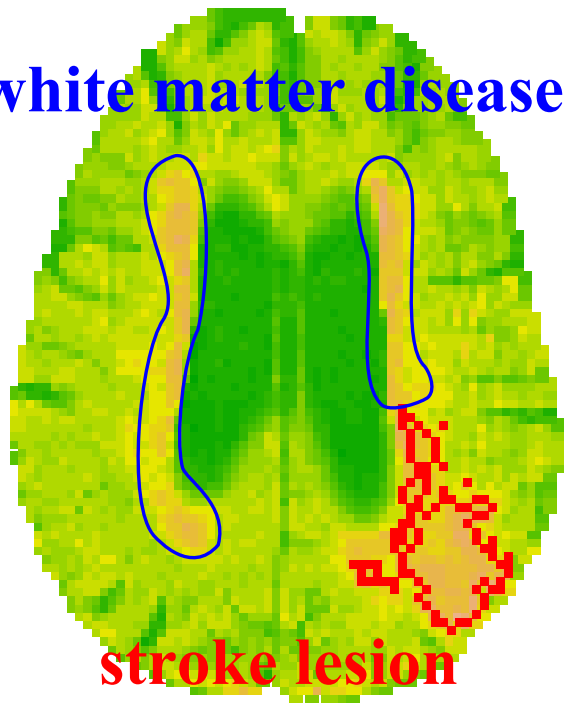
State of the art

Limits

- noise degrades the segmentation
  ▷ *univariate* spatial fMM
  (Woolrich et al., 2005 ; Feng, Tierney, and Magnotta, 2012 ; Zhang et al., 2008)

- white matter disease can be confused with stroke lesion
  ▷ lead to segmentation errors
  ▷ volume over-estimation

$\Rightarrow$ **need for a regional approach**



Objective

Propose an **unsupervised lesion segmentation algorithm robust** to noise and artefacts :

- allowing multivariate characterization of the lesion
- with a spatial regularization step :
  - ▷ local regularization for noise
  - ▷ regional regularization for artefacts

fMM - General Framework

Markov Random field : n sites where we observe an intensity Y

Mixture assumption : the observed intensity is issued from a mixture of G groups :

$$\mathbb{P}\left[Y_i|\Theta\right] = \sum_{g=1}^{G} \mathbb{P}\left[Y_i|\xi_i = g, \theta_g\right] \mathbb{P}\left[\xi_i = g\right]$$

with

- ξ_i : group membership of observation i
- θ_g : the distribution parameters of group g
- $\mathbb{P}\left[\xi_i = g\right]$: prior group membership of i for group g

Mean Field Approximation (MFA)

$$\mathbb{P}[Y|\Theta] = \prod_{i=1}^{n} ?? \text{ (no more independance)}$$

Mean field approximation : the neighboring group memberships are fixed to their expectation $\overline{\xi_{\mathcal{V}(i)}}$ (Zhang, 1992) :

$$\mathbb{P}^{MFA}[\xi] = \prod_{i=1}^{n} \mathbb{P}\left[\xi_i | \overline{\xi_{\mathcal{V}(i)}}\right] \approx \mathbb{P}[\xi]$$

One can show that the likelihood becomes :

$$\mathbb{P}[Y|\Theta] = \prod_{i=1}^{n} \sum_{g=1}^{G} \mathbb{P}[Y_i|\xi_i = g, \theta_g] \mathbb{P}\left[\xi_i = g | \overline{\xi_{\mathcal{V}(i)}}, \rho\right]$$

Probability distribution on a MRF

Hammersley-Clifford theorem

The join probability of a Markov Random Field (MRF) is a Gibbs distribution :

$$\mathbb{P}[\xi = x] = \frac{1}{Z} \exp[(\rho * U)(x)]$$

- U : spatial potential
- Z : normalizing constant

We define the spatial potential as the sum of :

- a **local potential** U_{loc} with intensity ρ_1
- a **regional potential** U_{reg} with intensity ρ_2

$$\mathbb{P}\left[\xi_i = g|\overline{\xi_{\mathcal{V}(i)}}\right] = \frac{1}{Z_i} \exp\left[\rho_1 U_{loc,g}(\overline{\xi_{\mathcal{V}(i,1)}}) + \rho_2 U_{reg,g}(\overline{\xi_{\mathcal{V}(i,C)}})\right]$$

Standard Potts model

- define the local neighborhood $\mathcal{V}(i,1)$

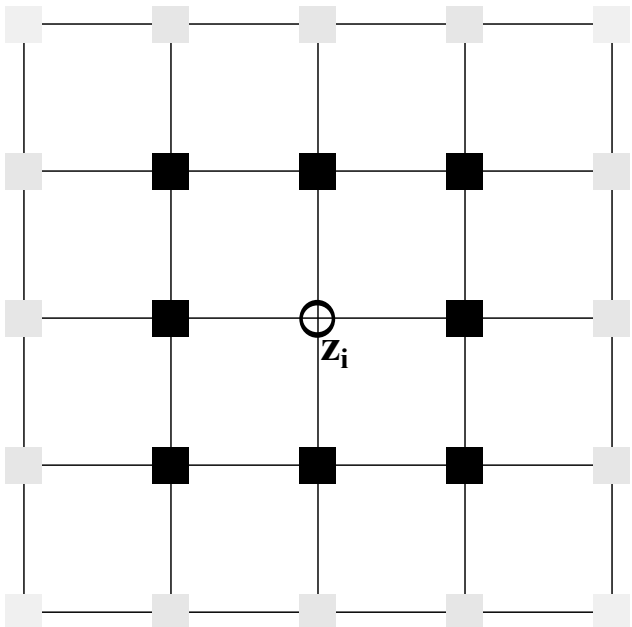


Figure: Queen's neighborhood

- compute the local potential

$$U_{loc,g}(\overline{\xi_{\mathcal{V}(i,1)}}) = \frac{1}{card\mathcal{V}(i,1)} \sum_{j \in \mathcal{V}(i,1)} \mathbb{P}\left[\overline{\xi_j} = g\right]$$

Regional Potts model

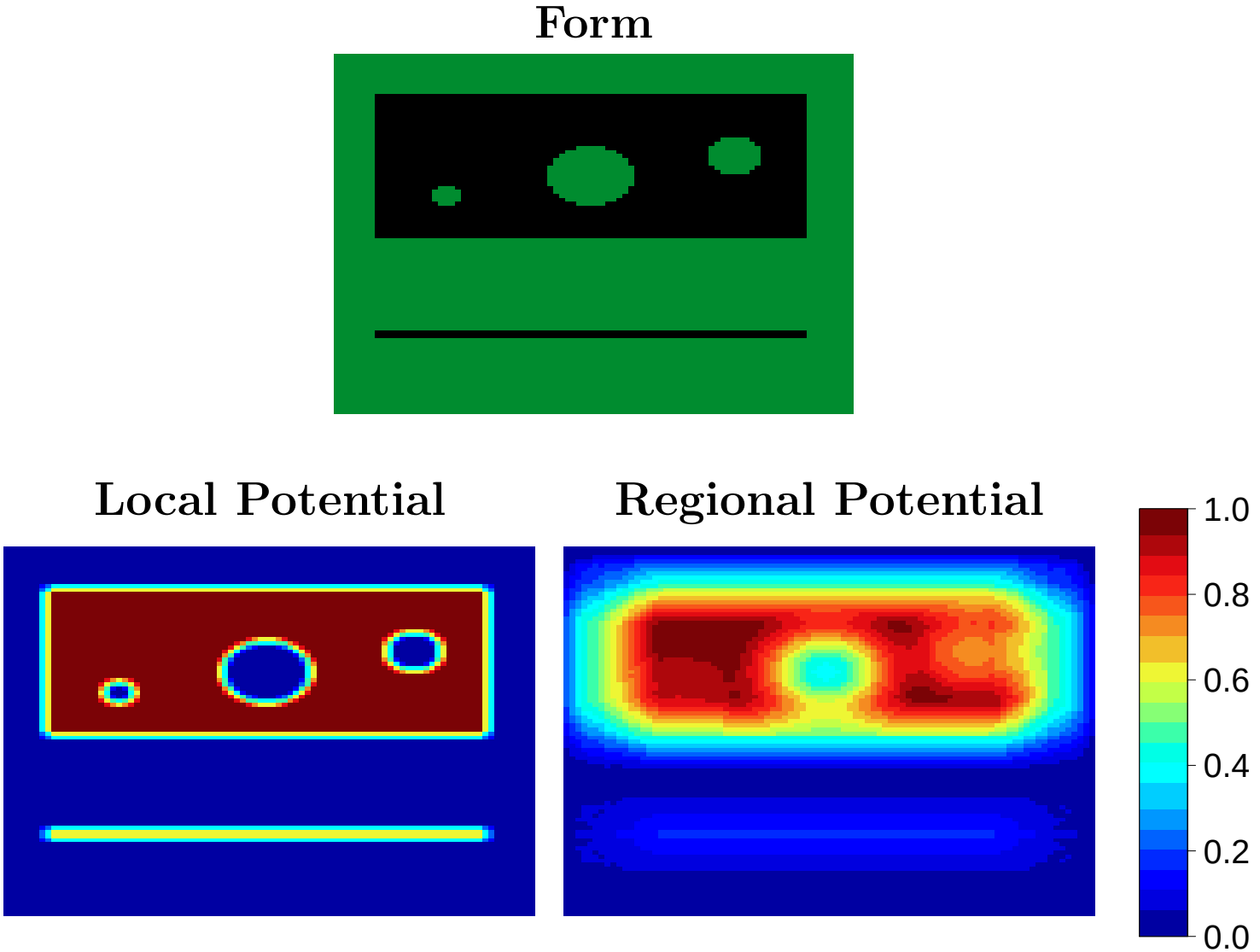
- define a regional neighbourhood as C neighbourhoods with **increasing range** $\mathcal{V}(i,c)$.
- compute the potential for each neighbourhood :

$$U_g^c(\overline{\xi_{\mathcal{V}(i,c)}}) = \frac{1}{card\mathcal{V}(i,c)} \sum_{j \in \mathcal{V}(i,c)} \mathbb{P}\left[\overline{\xi_j} = g\right]$$

- the regional potential is the average of these potentials

$$U_{reg}(\overline{\xi_{\mathcal{V}(i,c)}}) = \frac{1}{C} \sum_{c=1}^{C} U_g^c(\overline{\xi_{\mathcal{V}(i,c)}})$$

Local vs Regional Potential



Estimation - EM algorithm

- *step E* : estimate the group membership $\xi_{i,g}$
  - ▷ initialize the membership probabilities : $\mathbb{P}\left[\xi_i = g | Y_i\right]$
  - ▷ estimate the regularized membership probabilities iteratively over sites : $\mathbb{P}\left[\xi_i = g | \overline{\xi_{\mathcal{V}(i)}}, \rho\right]$

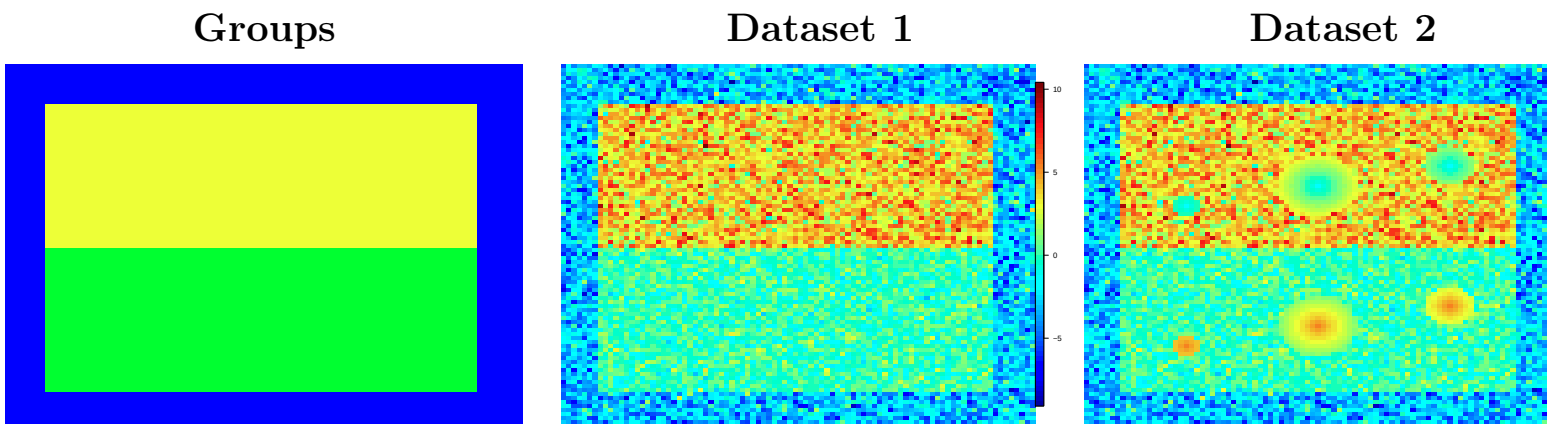
- *step M* : optimize the distribution Θ parameters

Simulation setup

Scenari

- **Dataset 1*** : 3 groups following normal laws with :
  mean : $\mu = (-3, 0, 4)$
  variance : $\sigma^2 = (3, 1, 3)$
- **Dataset 2** : same as scenario 1 with circular artefact

$\Longrightarrow$ The objective is to identify group 3 ('yellow' group)



* same simulation as in Woolrich et al., 2005

Simulation setup

Scenari

- **Dataset 1*** : 3 groups following normal laws with :
  mean : $\mu = (-3, 0, 4)$
  variance : $\sigma^2 = (3, 1, 3)$
- **Dataset 2** : same as scenario 1 with circular artefact

$\Longrightarrow$ The objective is to identify group 3 ('yellow' group)

Model specification

3 models were compared :

- $\mathcal{M}_0$: $\rho_1 = 0$ and $\rho_2 = 0$
- $\mathcal{M}_{loc}$: $\rho_1 = 6$ and $\rho_2 = 0$
- $\mathcal{M}_{reg}$: $\rho_1 = 0$ and $\rho_2 = 6$

* same simulation as in Woolrich et al., 2005

Simulation results

Non spatial fMM :

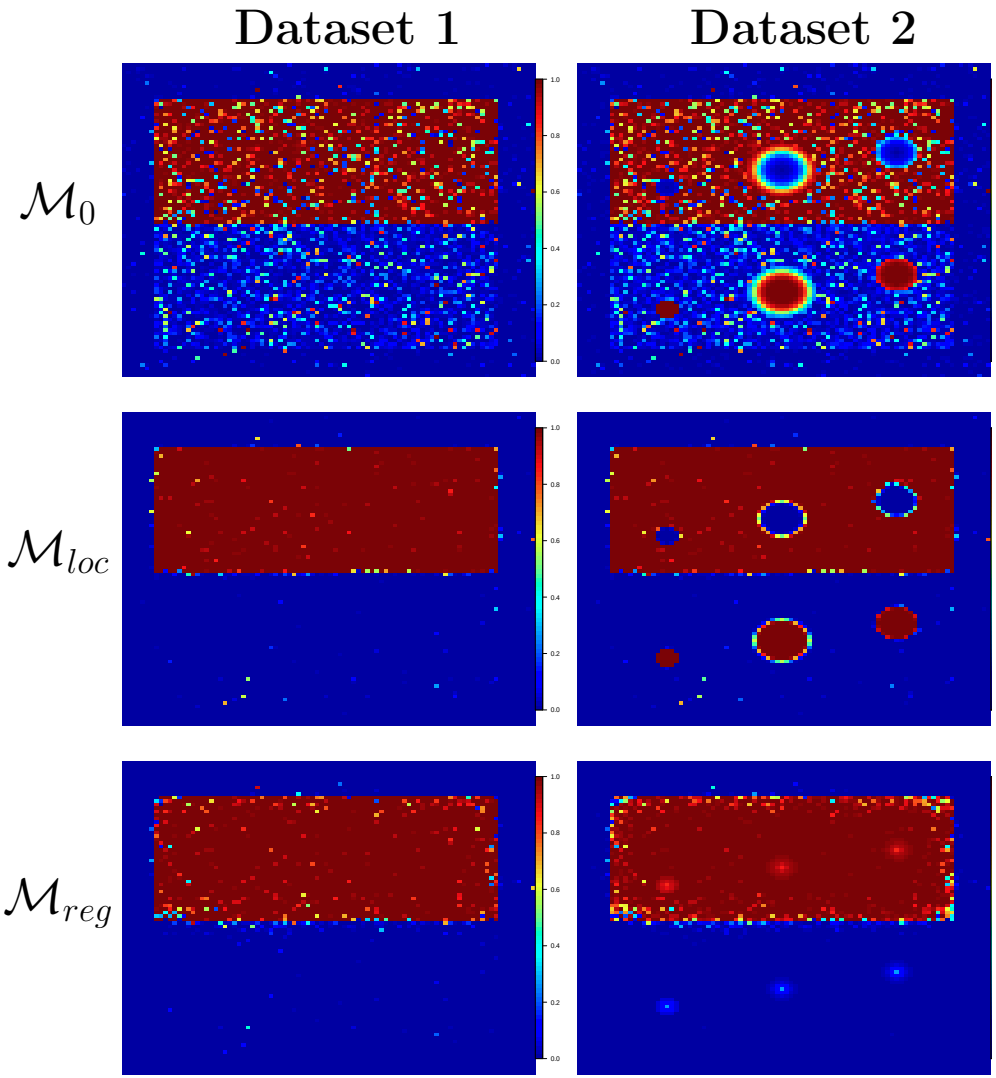
- noise and artefacts.

Local regularisation :

- noise correction.

Regional regularisation :

- artefacts correction.
- noise correction with edge effects.

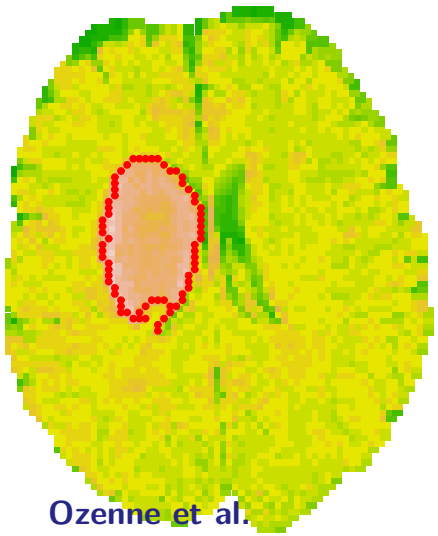


Stroke Segmentation

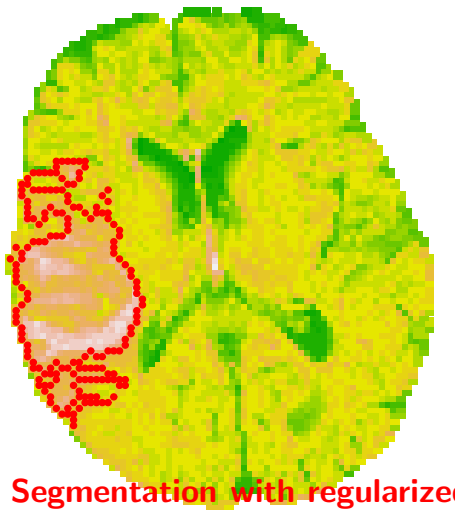
MRI data

- 9 patients with ischemic stroke from the I-know cohort
  - 4 'Typical'
  - 2 with 'Heterogeneity'
  - 3 with 'White Matter Disease'
- T2 FLAIR image at 1-month follow up
- physician segmentation (reference)

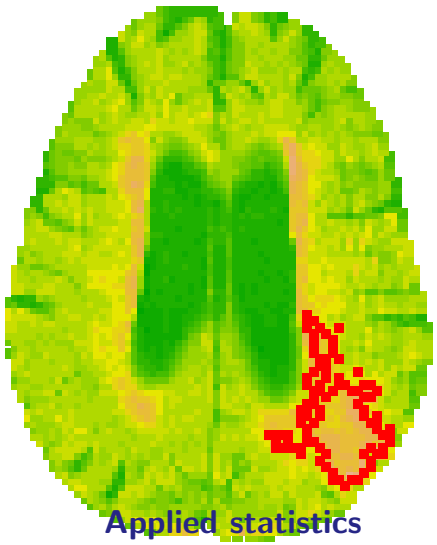
Typical



Heterogeneity



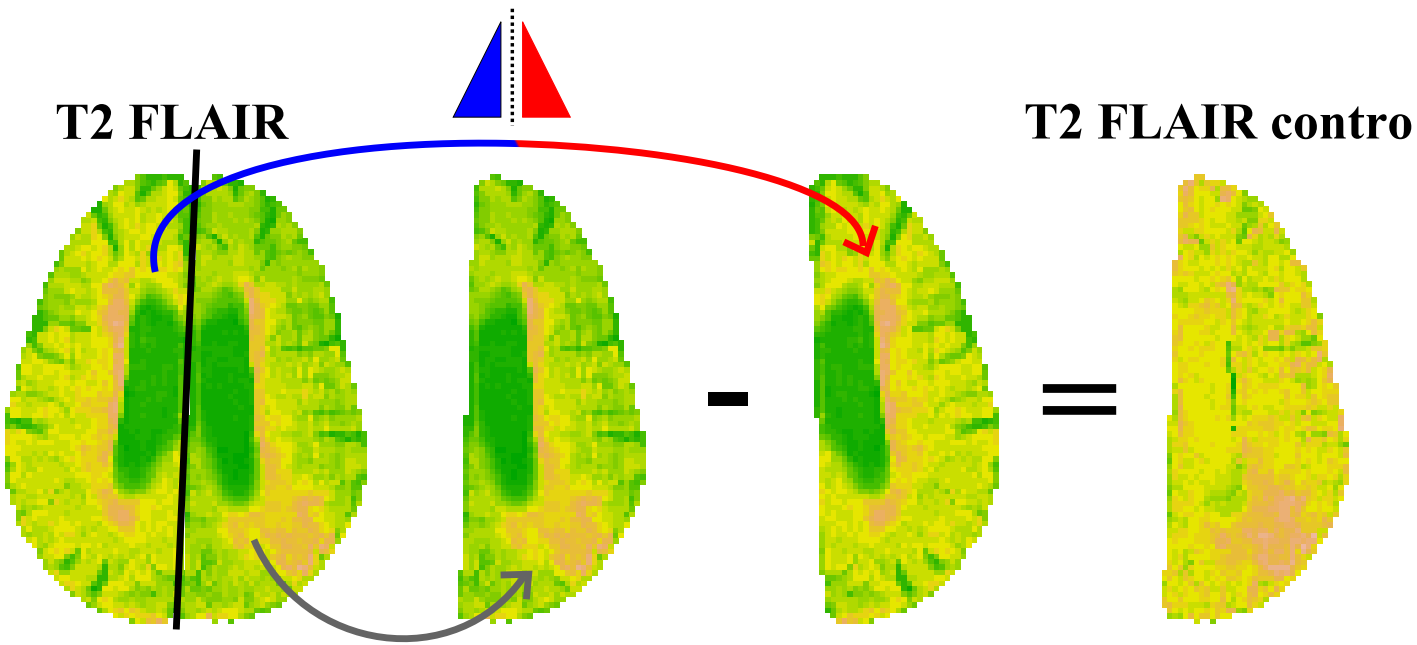
White Matter disease



Stroke Segmentation

main fMM settings

- 4 groups
- 2 parameters : T2 FLAIR and T2 FLAIR contro
- spatial parameters estimated on 'Typical' patients



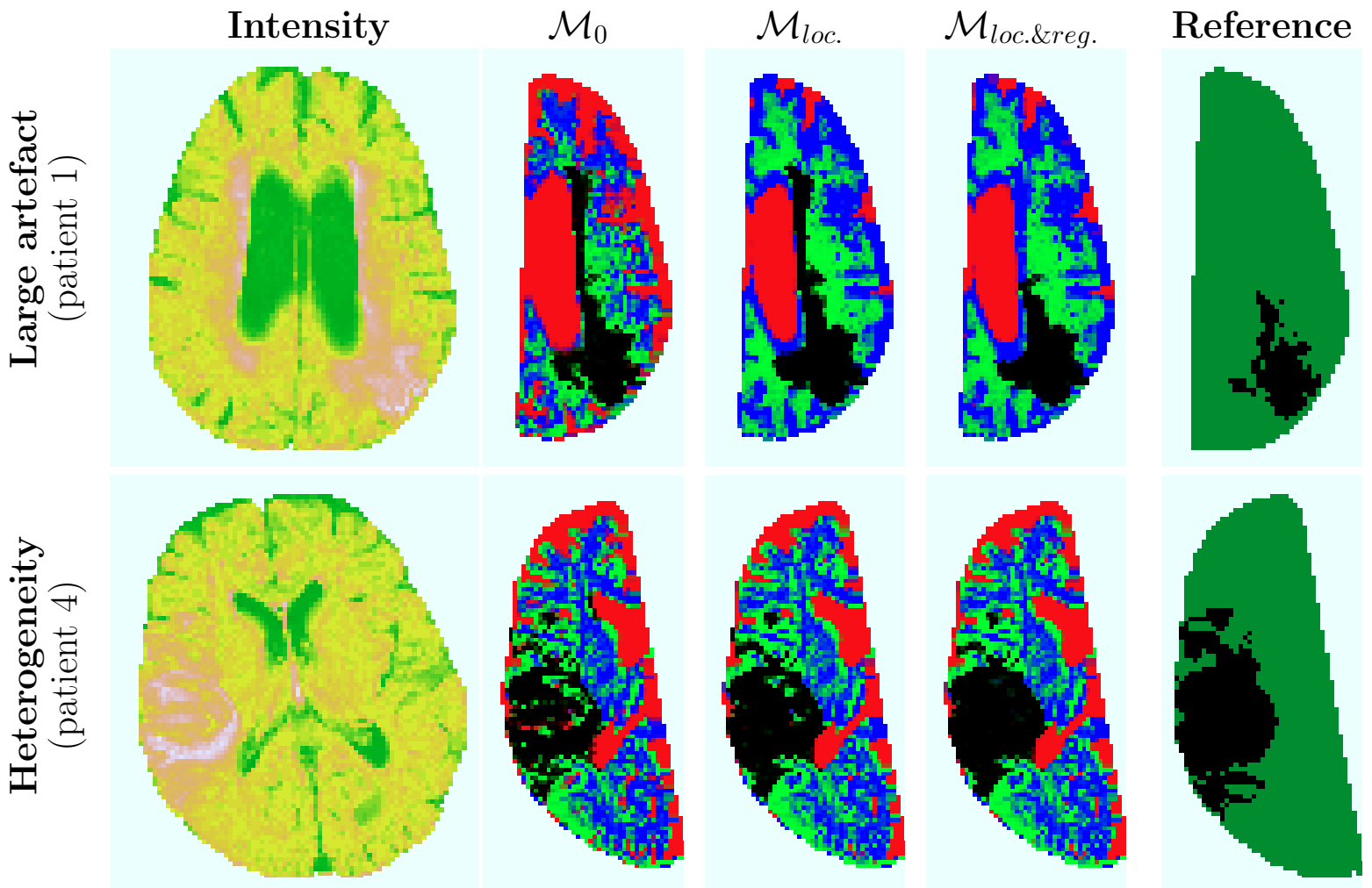
Stroke Segmentation - Results

Quality of the estimated volume (**1 is the optimum**) :

$$Quality = \frac{V_{model}}{V_{reference}}$$

		$\mathcal{M}_0$	$\mathcal{M}_{loc.}$	$\mathcal{M}_{loc.\&reg}$
White Matter disease	patient 1	1.62	1.68	1.17
	patient 2	2.76	1.16	1.10
	patient 3	4.85	1.41	1.29
Heterogeneity	patient 4	0.879	0.933	0.932
	patient 5	0.935	0.975	0.975

Stroke Segmentation - Results



Conclusion

Discussion

- Spatial regularization improves lesion segmentation :
  - ▷ local regularization deals with noise and heterogeneity.
  - ▷ regional regularization corrects artefacts (at least partially).
- A good initialisation is required to find optimal convergence.
  - ▷ k.means or non spatial fMM results.
- Estimation of the spatial parameters is still an issue.
  - ▷ automatic procedure is possible
    but underestimate the regional regularization parameter ρ_2

Perspectives

- Integration of the functions into a  package.
- Validation on a larger sample and with other diseases.

Bibliographie I

Feng, D., L. Tierney, and V. Magnotta (2012). "MRI Tissue Classification Using High-Resolution Bayesian Hidden Markov Normal Mixture Models". In: *Journal of the American Statistical Association* 107.497, pp. 102–119.

Klëppel, S. et al. (2011). "A comparison of different automated methods for the detection of white matter lesions in MRI data". In: *NeuroImage* 57.2, pp. 416–422. ISSN: 1053-8119.

Mouridsen, K. et al. (2013). "Acute Stroke: Automatic Perfusion Lesion Outlining Using Level Sets". In: *Radiology* 269.2. PMID: 23687176, pp. 404–412.

Osher, S. and R. Fedkiw (2003). *Level Set Methods and Dynamic Implicit Surfaces*. Springer.

Sweeney, E. M. et al. (2013). "OASIS is Automated Statistical Inference for Segmentation, with applications to multiple sclerosis lesion segmentation in MRI." In: *Neuroimage Clin.* 2, pp. 402–13.

Bibliographie II

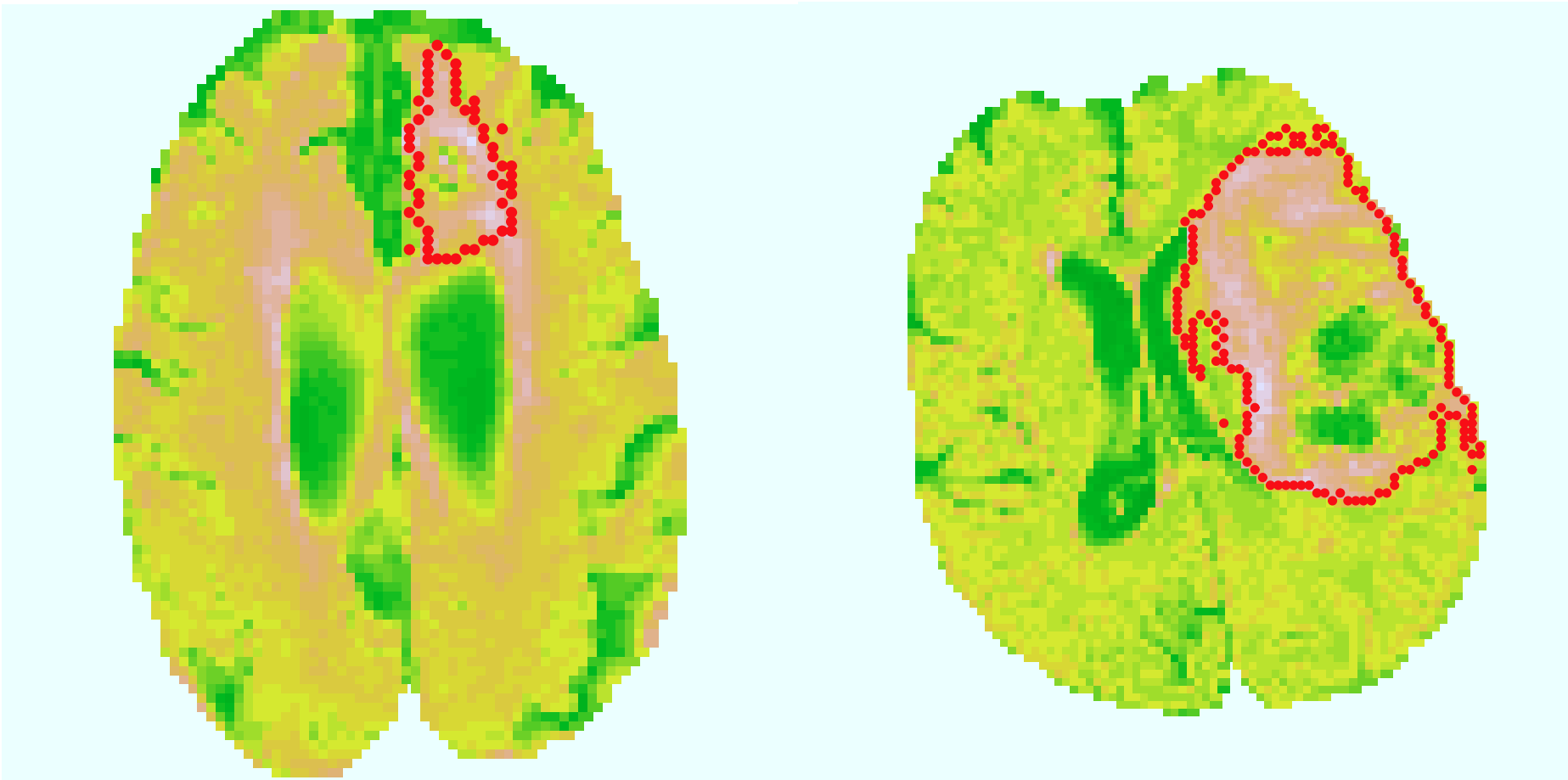
Weinman, J. et al. (2003). "Nonlinear Diffusion Scale-Space and Fast Marching Level Sets for Segmentation of MR Imagery and Volume Estimation of Stroke Lesions". In: *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*. Lecture Notes in Computer Science 2879. Ed. by R. Ellis and T. Peters, pp. 496–504.

Woolrich, M. et al. (2005). "Mixture models with adaptive spatial regularization for segmentation with an application to FMRI data". In: *IEEE Transactions on Medical Imaging* 24.1, pp. 1–11. ISSN: 0278-0062.

Zhang, J. (1992). "The mean field theory in EM procedures for Markov random fields". In: *Signal Processing, IEEE Transactions on* 40.10, pp. 2570–2583. ISSN: 1053-587X.

Zhang, X. et al. (2008). "Quantitative magnetic resonance image analysis via the EM algorithm with stochastic variation". In: *The Annals of Applied Statistics* 2.2, pp. 736–755.

Exemple of excluded patients



Likelihood for spatial fMM

$$
\begin{aligned}
L_v(\Theta|Y,X) &= \mathbb{P}[Y|\Theta,X] \\
&= \sum_{\Gamma=(g_1,\dots g_n)\in[1;G]^n} \mathbb{P}[Y,\xi=\Gamma|\Theta] \\
&\approx \sum_{\Gamma=(g_1,\dots g_n)\in[1;G]^n} \prod_{i=1}^{n} \mathbb{P}\left[Y_i,\xi_i=g_i|\Theta,\overline{\xi_{\mathcal{V}(i)}}\right] \\
&= L_v^{MFA}(\Theta|Y)
\end{aligned}
$$

using mean field approximation

Likelihood for spatial fMM

Then

$$L_v^{MFA}(\Theta|Y,X) = \sum_{\Gamma=(g_1,\dots g_n)\in[1;G]^n} \prod_{i=1}^{n} \mathbb{P}\left[Y_i, \xi_i = g_i | \Theta, \overline{\xi_{\mathcal{V}(i)}}\right]$$

$$= \sum_{\Gamma=(g_1,\dots g_n)\in[1;G]^n} \prod_{i=1}^{n} \mathbb{P}\left[Y_i|\xi_i = g_i, \theta_{g_i}\right] \mathbb{P}\left[\xi_i = g_i | \rho, \overline{\xi_{\mathcal{V}(i)}}\right]$$

$$= \prod_{i=1}^{n}\sum_{g=1}^{G} \mathbb{P}\left[Y_i|\xi_i = g, \theta_g\right] \mathbb{P}\left[\xi_i = g | \rho, \overline{\xi_{\mathcal{V}(i)}}\right]$$

$$= \prod_{i=1}^{n}\sum_{g=1}^{G} \mathbb{P}\left[Y_i|\xi_{i,g} = 1, \theta_g\right]$$

$$\times \frac{1}{Z} \exp\left[\rho_1 U_{loc}(\xi_{\mathcal{V}(i),g}) + \rho_2 U_{reg}(\xi_{\mathcal{V}(i),g})\right]$$

Complete likelihood for spatial fMM

Denoting : $\pi_{i,g}^{posterior} = \mathbb{P}\left[\xi_{i,g} | Y_i, \overline{\xi_{\mathcal{V}(i)}}, \Theta\right]$

$$L_v^c(Y_i|\Theta,\rho) = \prod_{i=1}^{n}\prod_{g=1}^{G}\left(\mathbb{P}\left[Y_i|\xi_{i,g}=1,\theta_g\right]\mathbb{P}\left[\xi_{i,g}=1|\overline{\xi_{\mathcal{V}(i)}},\rho\right]\right)^{\pi_{i,g}^{posterior}}$$

$$l_v^c(Y_i|\Theta,\rho) = \sum_{i=1}^{n}\sum_{g=1}^{G}\pi_{i,g}^{posterior}\log\mathbb{P}\left[Y_i|\xi_{i,g}=1,\theta_g\right] + \pi_{i,g}^{posterior}\log\mathbb{P}\left[\xi_{i,g}=1|\overline{\xi_{\mathcal{V}(i)}},\rho\right]$$

- $\mathcal{M}_{intensity}$: sum of independent weighted glm models
  $\Rightarrow$ IWLS
- $\mathcal{M}_{spatial}$: local and regional Potts model
  $\Rightarrow$ quasi-Newton method (L-BFGS-B)

Estimation of the spatial parameter - Method

$$
\begin{aligned}
& \mathcal{M}_{spatial}(\rho) \\
= & \sum_{i=1}^{n}\sum_{g=1}^{G} \pi_{i,g}^{posterior} \log \frac{1}{Z_i} \exp\left[\rho_1 U_{loc}(\overline{\xi_{\mathcal{V}(i),g}}) + \rho_2 U_{reg}(\overline{\xi_{\mathcal{V}(i),g}})\right] \\
= & \sum_{i=1}^{n}\sum_{g=1}^{G} \pi_{i,g}^{posterior} \left(-\log Z_i + \rho_1 U_{loc}(\overline{\xi_{\mathcal{V}(i),g}}) + \rho_2 U_{reg}(\overline{\xi_{\mathcal{V}(i),g}})\right)
\end{aligned}
$$

with $Z_i = \sum_{g=1}^{G} \exp\left[\rho_1 U_{loc}(\overline{\xi_{\mathcal{V}(i),g}}) + \rho_2 U_{reg}(\overline{\xi_{\mathcal{V}(i),g}})\right]$

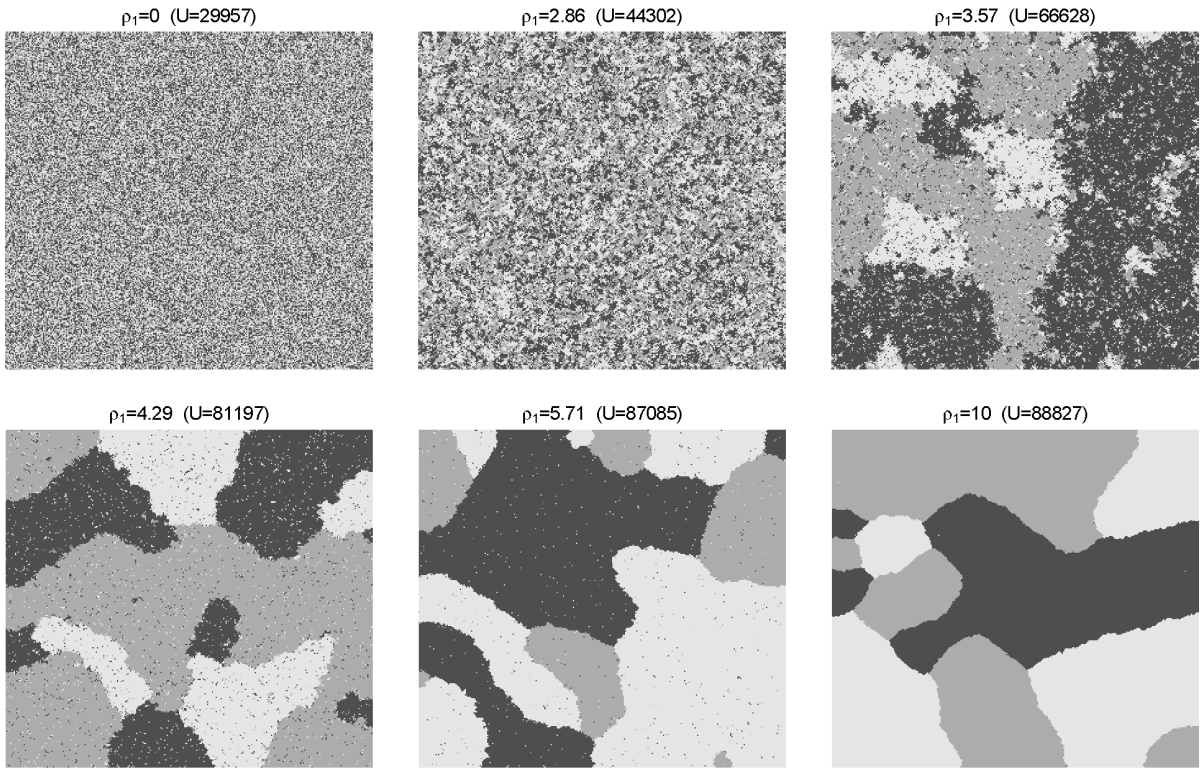
U_{loc} and U_{reg} can be computed for each patient and thus also Z_i.
The function to optimize is a two parameter function that is derivable
$\Rightarrow$ quasi-Newton method.

Estimation of the spatial parameter - Results

type	median estimation	range
local reg. (ρ_1)	4.59	$[4.32 - 5.17]$
local and regional reg. (ρ_1)	3.85	$[3.47 - 4.77]$
local and regional reg. (ρ_2)	2.61	$[0.20 - 4.08]$

Validity of MFA - Simulation

- Potts model simulated by Gibbs sampling (1000 iterations).
- n ranged from 100 to 1000.
- ρ_1 ranged from 0 to 10.
- each scenario was replicated 250 times.



Validity of MFA - Results

- clear decrease in variance when n increases.
- small decrease in bias when n increases.
- MRI data (n $\sim$ 30000) : relative bias $<$5 % for common ρ_1

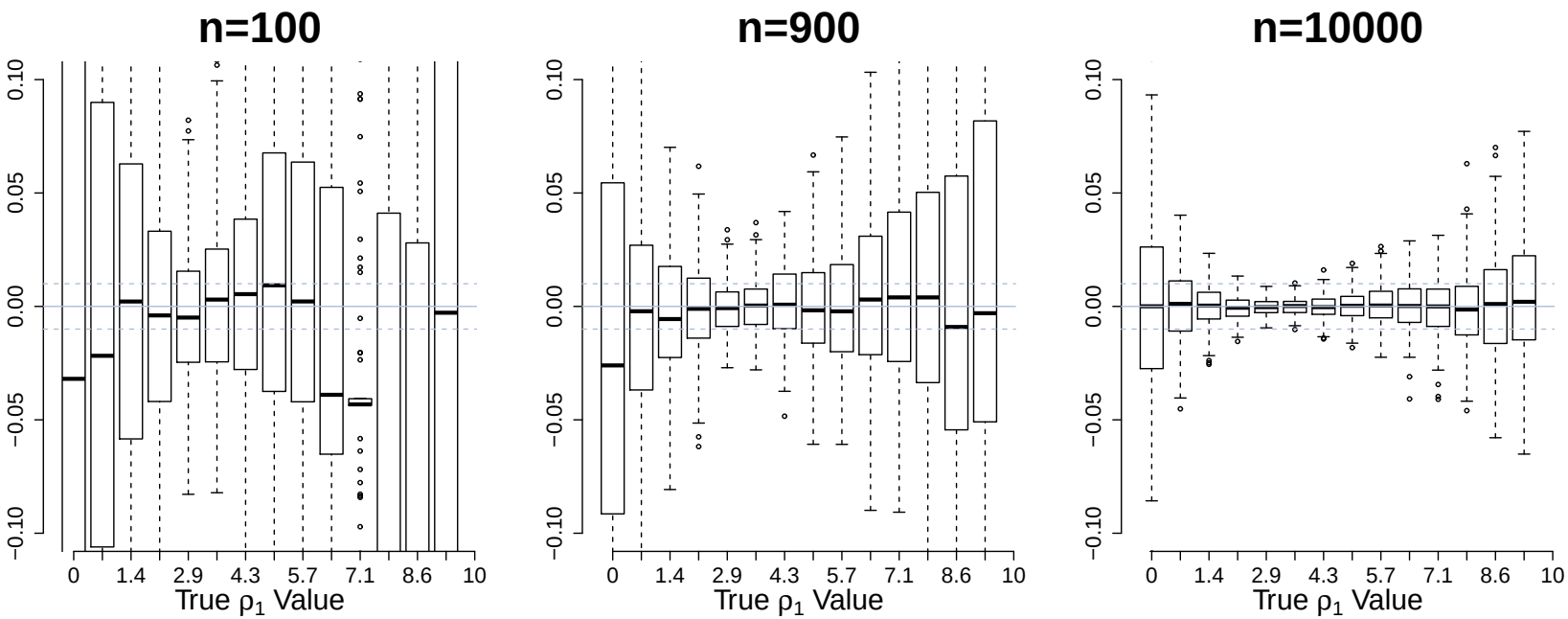


Figure: Relative bias of the ρ_1 estimator

fMMseg - Example

```
> require(fMMseg)
> data(data_test,package="fMMseg")
> str(data_test)

'data.frame': 8100 obs. of  6 variables:
 $ i              : int  1 2 3 4 5 6 7 8 9 10 ...
 $ j              : int  1 1 1 1 1 1 1 1 1 1 ...
 $ group          : num  1 1 1 1 1 1 1 1 1 1 ...
 $ Y              : num  -2.97 -3.32 -5.38 -4.04 -2.49 ...
 $ Y_artlinear    : num  -2.97 -3.32 -5.38 -4.04 -2.49 ...
 $ Y_artspherique: num  -2.97 -3.32 -5.38 -4.04 -2.49 ...
```

fMMseg - Example

```
> res_test <- EM.launcher(G=3,data=data_test,coords=c("i","j"),
+                         distband.SR=sqrt(2),distband.LR=10,
+      var_reg="Y_artlinear",family=gaussian(link="identity"),
+                          test.ICM=T,rho_ICM=c(6,6),
+                          test.ICMregional=T)

* initialisation by k means *
 # initialisation
                groupe 1   groupe 2   groupe 3
intercept 1 :  -3.702351 0.02097692   4.384099
cv criteria :  0.001
*** init. spatial regularization ***
### Iteration FINALE  20  (lv = -15521.41)
                groupe 1   groupe 2   groupe 3
intercept 1 : -3.0083412  0.1448128  3.8479213
sigma 1 :      1.7505830  1.1915927  1.8669002
<prior> :      0.3333333  0.3333333  0.3333333
<posterior> :  0.3330471  0.3308790  0.3360739
ICM parameters :  6 6
*** Convergence ***
*** Export ***
```