

# Inverse Problems in Signal and Image processing, Imaging systems and Computer Vision

From Deterministic Regularization to  
Probabilistic Bayesian Approaches

Ali Mohammad-Djafari

Groupe Problèmes Inverses  
Laboratoire des signaux et systèmes (L2S)  
UMR 8506 CNRS - SUPELEC - UNIV PARIS SUD 11  
Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette, FRANCE.

[djafari@lss.supelec.fr](mailto:djafari@lss.supelec.fr)  
<http://djafari.free.fr>  
<http://www.lss.supelec.fr>

Master ITEMS 2012, Université Polytechnique de Bucarest,

Nov 2012

# Content

- ▶ Invers problems : Examples and general formulation
- ▶ Inversion methods :  
analytical, parametric and non parametric
- ▶ Deterministic methods:  
Data matching, Least Squares, Regularization
- ▶ Probabilistic methods:  
Probability matching, Maximum likelihood, Bayesian inference
- ▶ Bayesian inference approach
- ▶ Prior models for images
- ▶ Bayesian computation
- ▶ Applications:  
Computed Tomography, Image separation, Superresolution,  
SAR Imaging
- ▶ Conclusions
- ▶ Questions and Discussion

# Inverse problems : 3 main examples

- ▶ Example 1:  
Measuring variation of temperature with a thermometer
  - ▶  $f(t)$  variation of temperature over time
  - ▶  $g(t)$  variation of length of the liquid in thermometer
- ▶ Example 2:  
Making an image with a camera, a microscope or a telescope
  - ▶  $f(x, y)$  real scene
  - ▶  $g(x, y)$  observed image
- ▶ Example 3: Making an image of the interior of a body
  - ▶  $f(x, y)$  a section of a real 3D body  $f(x, y, z)$
  - ▶  $g_\phi(r)$  a line of observed radiograph  $g_\phi(r, z)$
- ▶ Example 1: Deconvolution
- ▶ Example 2: Image restoration
- ▶ Example 3: Image reconstruction

# Measuring variation of temperature with a thermometer

- ▶  $f(t)$  variation of temperature over time
- ▶  $g(t)$  variation of length of the liquid in thermometer
- ▶ Forward model: Convolution

$$g(t) = \int f(t') h(t - t') dt' + \epsilon(t)$$

$h(t)$ : impulse response of the measurement system

- ▶ Inverse problem: Deconvolution

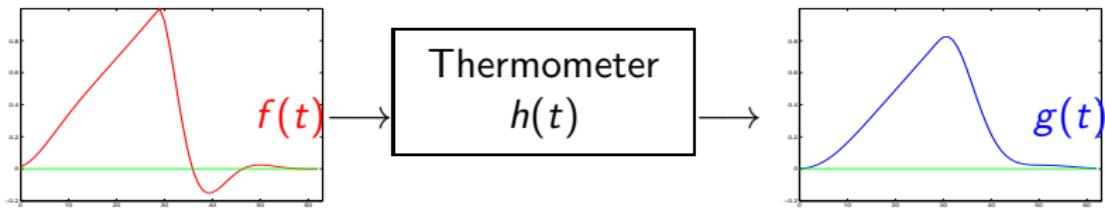
Given the forward model  $\mathcal{H}$  (impulse response  $h(t)$ ))  
and a set of data  $g(t_i), i = 1, \dots, M$   
find  $f(t)$



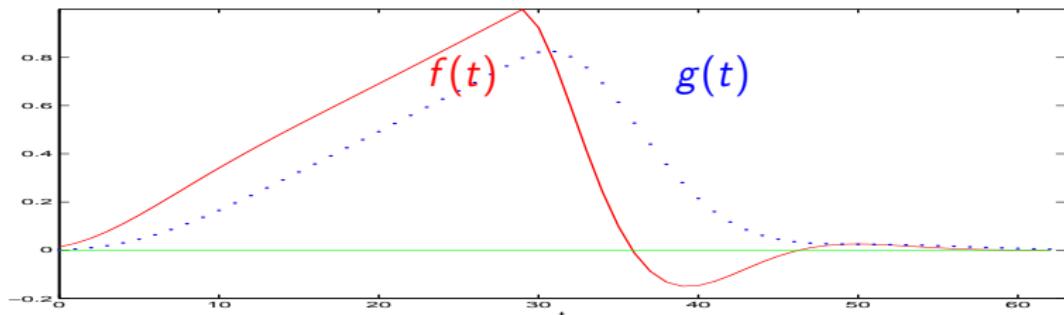
# Measuring variation of temperature with a thermometer

Forward model: Convolution

$$g(t) = \int f(t') h(t - t') dt' + \epsilon(t)$$

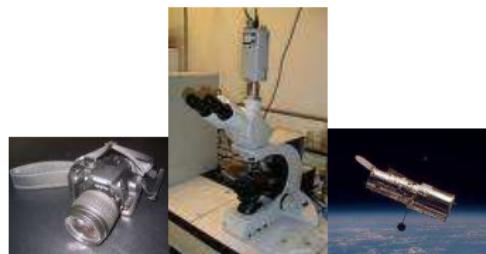


Inversion: Deconvolution



# Making an image with a camera, a microscope or a telescope

- ▶  $f(x, y)$  real scene
- ▶  $g(x, y)$  observed image
- ▶ Forward model: Convolution



$$g(x, y) = \iint f(x', y') h(x - x', y - y') dx' dy' + \epsilon(x, y)$$

$h(x, y)$ : Point Spread Function (PSF) of the imaging system

- ▶ Inverse problem: Image restoration

Given the forward model  $\mathcal{H}$  (PSF  $h(x, y)$ ))

and a set of data  $g(x_i, y_i), i = 1, \dots, M$

find  $f(x, y)$

# Making an image with an unfocused camera

Forward model: 2D Convolution

$$g(x, y) = \iint f(x', y') h(x - x', y - y') dx' dy' + \epsilon(x, y)$$



$$f(x, y)$$

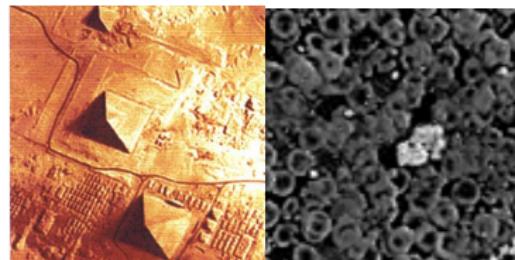
$$h(x, y)$$

$$\epsilon(x, y)$$

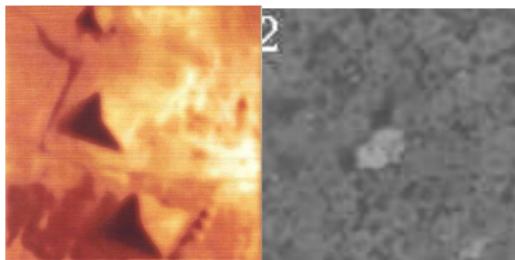


$$g(x, y)$$

Inversion: Deconvolution

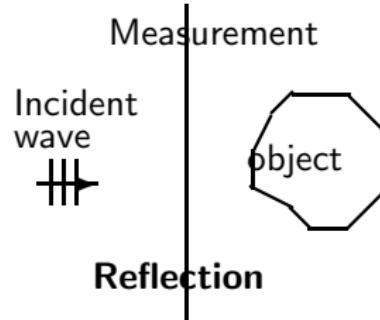
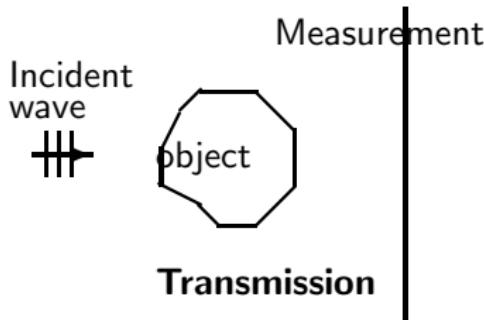
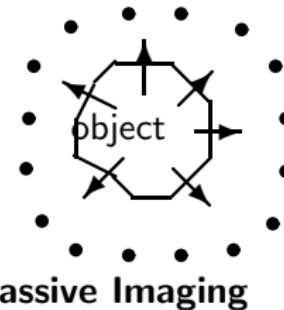
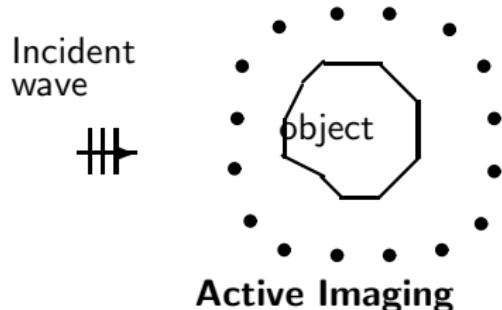


$$?$$
  
$$\Leftarrow$$



# Making an image of the interior of a body

## Different imaging systems:



**Forward problem:** Knowing the **object** predict the **data**

**Inverse problem:** From **measured data** find the **object**

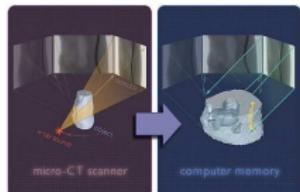
# Making an image of the interior of a body

- ▶  $f(x, y)$  a section of a real 3D body  $f(x, y, z)$
- ▶  $g_\phi(r)$  a line of observed radiograph  $g_\phi(r, z)$
- ▶ Forward model:  
Line integrals or Radon Transform

$$\begin{aligned} g_\phi(r) &= \int_{L_{r,\phi}} f(x, y) \, dl + \epsilon_\phi(r) \\ &= \iint f(x, y) \delta(r - x \cos \phi - y \sin \phi) \, dx \, dy + \epsilon_\phi(r) \end{aligned}$$

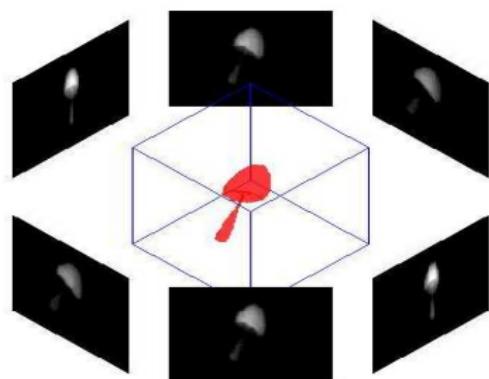
- ▶ Inverse problem: Image reconstruction

Given the forward model  $\mathcal{H}$  (Radon Transform) and  
a set of data  $g_{\phi_i}(r), i = 1, \dots, M$   
find  $f(x, y)$

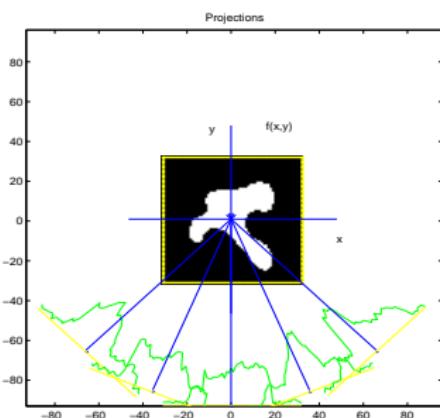


# 2D and 3D Computed Tomography

3D



2D



$$g_\phi(r_1, r_2) = \int_{\mathcal{L}_{r_1, r_2, \phi}} f(x, y, z) \, dl \quad g_\phi(r) = \int_{\mathcal{L}_{r, \phi}} f(x, y) \, dl$$

Forward problem:  $f(x, y)$  or  $f(x, y, z)$   $\rightarrow$   $g_\phi(r)$  or  $g_\phi(r_1, r_2)$

Inverse problem:  $g_\phi(r)$  or  $g_\phi(r_1, r_2)$   $\rightarrow$   $f(x, y)$  or  $f(x, y, z)$

# Microwave or ultrasound imaging

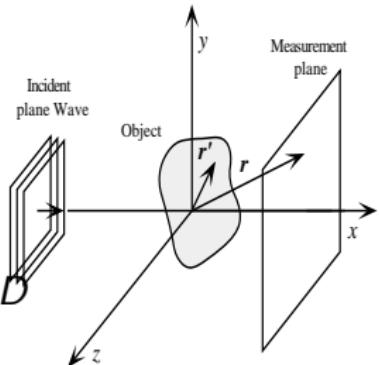
Measrs: diffracted wave by the object  $g(\mathbf{r}_i)$

Unknown quantity:  $f(\mathbf{r}) = k_0^2(n^2(\mathbf{r}) - 1)$

Intermediate quantity :  $\phi(\mathbf{r})$

$$g(\mathbf{r}_i) = \iint_D G_m(\mathbf{r}_i, \mathbf{r}') \phi(\mathbf{r}') f(\mathbf{r}') d\mathbf{r}', \quad \mathbf{r}_i \in S$$

$$\phi(\mathbf{r}) = \phi_0(\mathbf{r}) + \iint_D G_o(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}') f(\mathbf{r}') d\mathbf{r}', \quad \mathbf{r} \in D$$

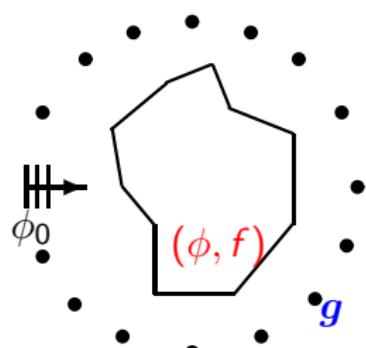


**Born approximation** ( $\phi(\mathbf{r}') \simeq \phi_0(\mathbf{r}')$ ):

$$g(\mathbf{r}_i) = \iint_D G_m(\mathbf{r}_i, \mathbf{r}') \phi_0(\mathbf{r}') f(\mathbf{r}') d\mathbf{r}', \quad \mathbf{r}_i \in S$$

**Discretization :**

$$\begin{cases} \mathbf{g} = \mathbf{G}_m \mathbf{F} \boldsymbol{\phi} \\ \boldsymbol{\phi} = \boldsymbol{\phi}_0 + \mathbf{G}_o \mathbf{F} \boldsymbol{\phi} \end{cases} \xrightarrow{\text{with}} \begin{cases} \mathbf{g} = \mathbf{H}(\mathbf{f}) \\ \text{with } \mathbf{F} = \text{diag}(\mathbf{f}) \\ \mathbf{H}(\mathbf{f}) = \mathbf{G}_m \mathbf{F} (\mathbf{I} - \mathbf{G}_o \mathbf{F})^{-1} \boldsymbol{\phi}_0 \end{cases}$$



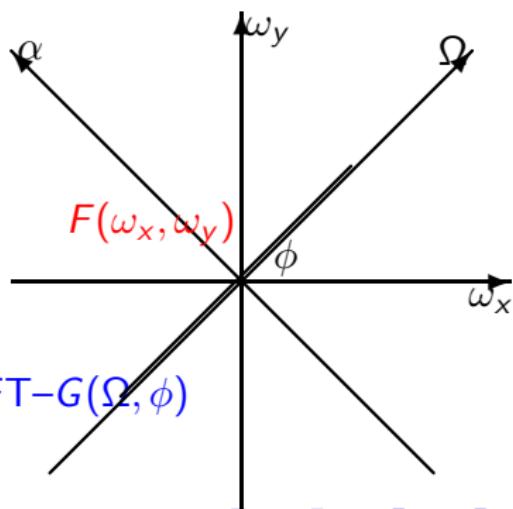
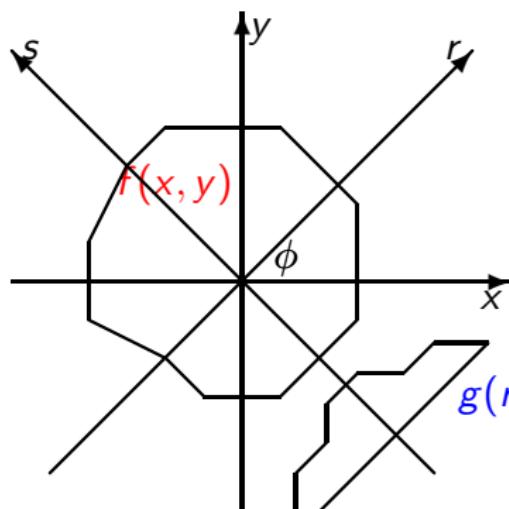
# Fourier Synthesis in X ray Tomography

$$g(r, \phi) = \iint f(x, y) \delta(r - x \cos \phi - y \sin \phi) dx dy$$

$$G(\Omega, \phi) = \int g(r, \phi) \exp \{-j\Omega r\} dr$$

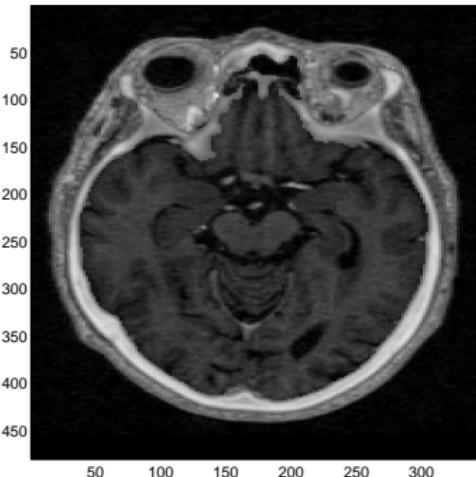
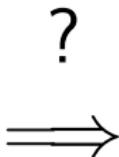
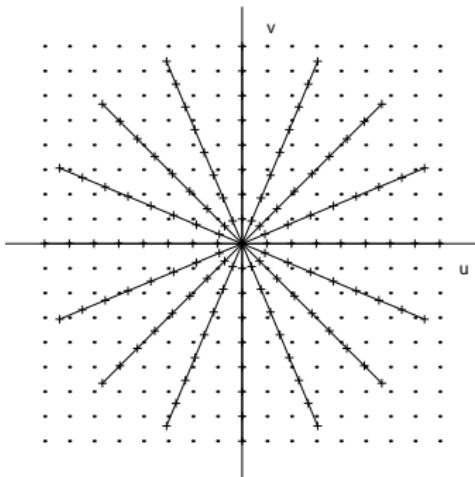
$$F(\omega_x, \omega_y) = \iint f(x, y) \exp \{-j\omega_x x, \omega_y y\} dx dy$$

$$F(\omega_x, \omega_y) = G(\Omega, \phi) \quad \text{for } \omega_x = \Omega \cos \phi \quad \text{and} \quad \omega_y = \Omega \sin \phi$$



# Fourier Synthesis in X ray tomography

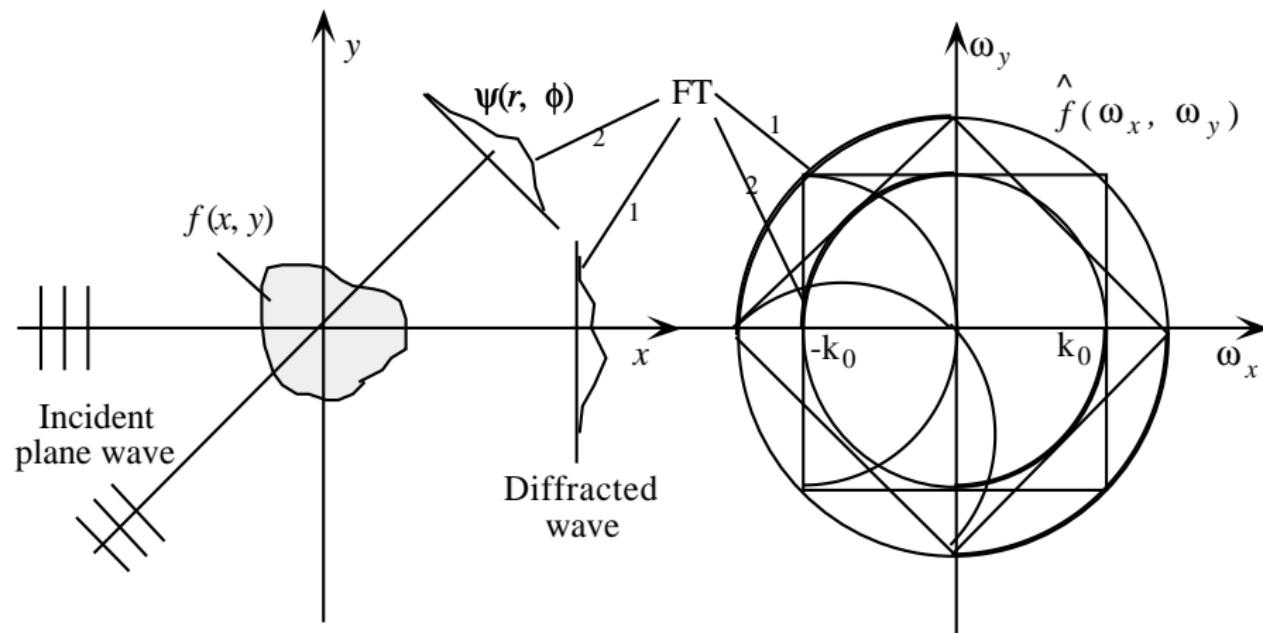
$$G(\omega_x, \omega_y) = \iint f(x, y) \exp \{-j(\omega_x x + \omega_y y)\} dx dy$$



**Forward problem:** Given  $f(x, y)$  compute  $G(\omega_x, \omega_y)$

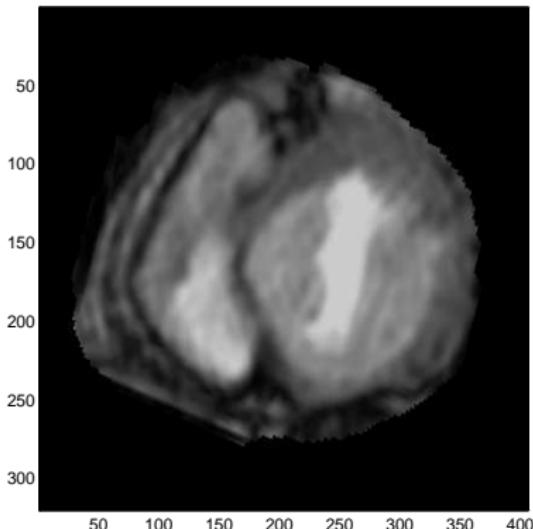
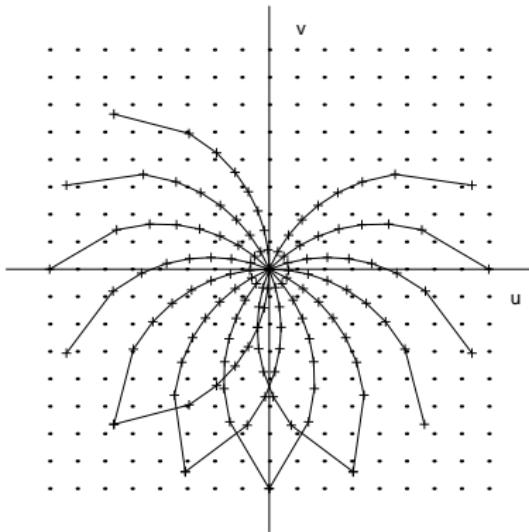
**Inverse problem:** Given  $G(\omega_x, \omega_y)$  on those lines  
estimate  $f(x, y)$

# Fourier Synthesis in Diffraction tomography



# Fourier Synthesis in Diffraction tomography

$$G(\omega_x, \omega_y) = \iint f(x, y) \exp \{-j(\omega_x x + \omega_y y)\} dx dy$$

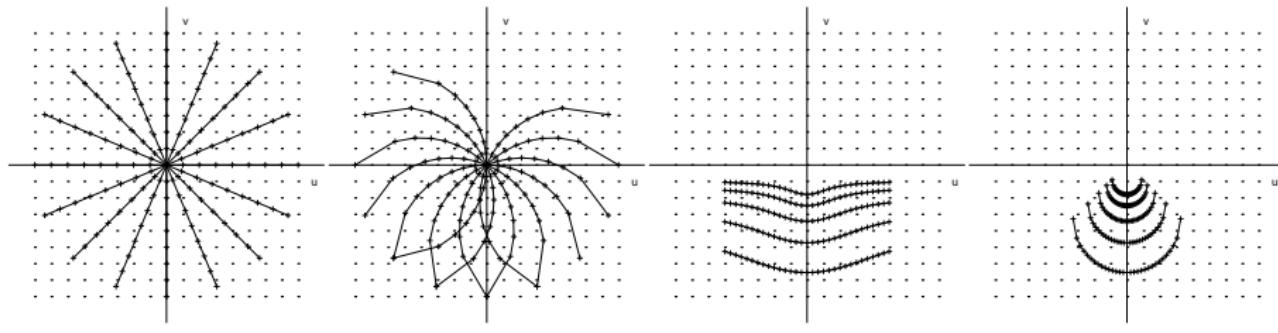


**Forward problem:** Given  $f(x, y)$  compute  $G(\omega_x, \omega_y)$

**Inverse problem :** Given  $G(\omega_x, \omega_y)$  on those semi circles  
estimate  $f(x, y)$

# Fourier Synthesis in different imaging systems

$$G(\omega_x, \omega_y) = \iint f(x, y) \exp \{-j(\omega_x x + \omega_y y)\} dx dy$$



X ray Tomography

Diffraction

Eddy current

SAR & Radar

**Forward problem:** Given  $f(x, y)$  compute  $G(\omega_x, \omega_y)$

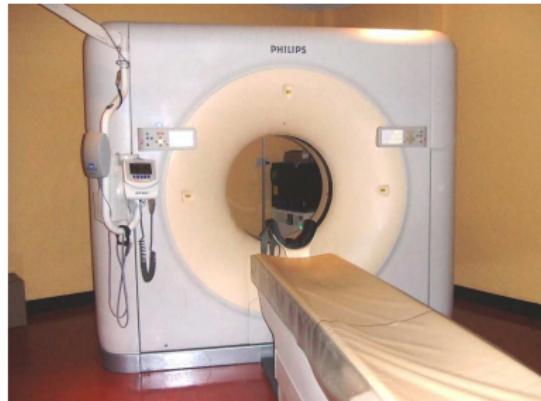
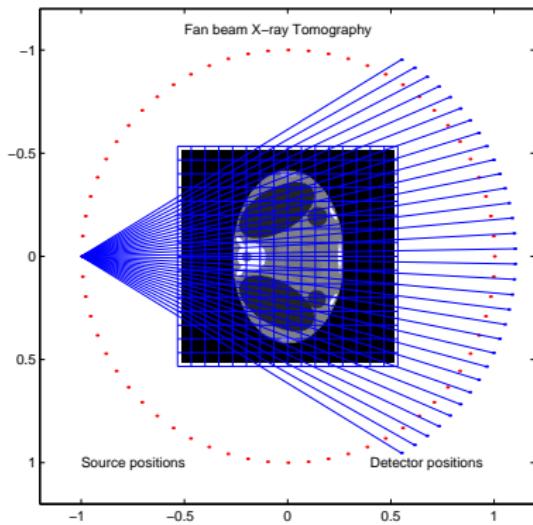
**Inverse problem :** Given  $G(\omega_x, \omega_y)$  on those algebraic lines, circles or curves, estimate  $f(x, y)$

# Invers Problems: other examples and applications

- ▶ X ray, Gamma ray Computed Tomography (CT)
- ▶ Microwave and ultrasound tomography
- ▶ Positron emission tomography (PET)
- ▶ Magnetic resonance imaging (MRI)
- ▶ Photoacoustic imaging
- ▶ Radio astronomy
- ▶ Geophysical imaging
- ▶ Non Destructive Evaluation (NDE) and Testing (NDT) techniques in industry
- ▶ Hyperspectral imaging
- ▶ Earth observation methods (Radar, SAR, IR, ...)
- ▶ Survey and tracking in security systems

# Computed tomography (CT)

## A Multislice CT Scanner



$$g(s_i) = \int_{L_i} f(r) \, dl_i + \epsilon(s_i)$$

Discretization

$$\mathbf{g} = \mathbf{Hf} + \boldsymbol{\epsilon}$$

# Positron emission tomography (PET)



# Magnetic resonance imaging (MRI)

Nuclear magnetic resonance imaging (NMRI), Para-sagittal MRI of the head



## Radio astronomy (interferometry imaging systems)

The Very Large Array in New Mexico, an example of a radio telescope.



# General formulation of inverse problems

- ▶ General non linear inverse problems:

$$g(s) = [\mathcal{H}f(r)](s) + \epsilon(s), \quad r \in \mathcal{R}, \quad s \in \mathcal{S}$$

- ▶ Linear models:

$$g(s) = \int f(r) h(r, s) dr + \epsilon(s)$$

If  $h(r, s) = h(r - s)$  → Convolution.

- ▶ Discrete data:

$$g(s_i) = \int h(s_i, r) f(r) dr + \epsilon(s_i), \quad i = 1, \dots, m$$

- ▶ Inversion: Given the forward model  $\mathcal{H}$  and the data

$g = \{g(s_i), i = 1, \dots, m\}$  estimate  $f(r)$

- ▶ Well-posed and **Ill-posed** problems (Hadamard):  
existence, uniqueness and stability

- ▶ Need for prior information

# Analytical methods (mathematical physics)

$$g(s_i) = \int h(s_i, r) \mathbf{f}(r) dr + \epsilon(s_i), \quad i = 1, \dots, m$$

$$g(s) = \int h(s, r) \mathbf{f}(r) dr$$

$$\widehat{\mathbf{f}}(r) = \int w(s, r) g(s) ds$$

$w(s, r)$  minimizing a criterion:

$$\begin{aligned} Q(w(s, r)) &= \|g(s) - [\mathcal{H}\widehat{\mathbf{f}}(r)](s)\|_2^2 = \int |g(s) - [\mathcal{H}\widehat{\mathbf{f}}(r)](s)|^2 ds \\ &= \int |g(s) - \int h(s, r) \widehat{\mathbf{f}}(r) dr|^2 ds \\ &= \int |g(s) - \int \int h(s, r) w(s, r) g(s) ds dr|^2 ds \end{aligned}$$

Trivial solution:  $h(s, r)w(s, r) = \delta(r)\delta(s)$

# Analytical methods

- Trivial solution:

$$w(s, r) = h^{-1}(s, r)$$

Example: Fourier Transform:

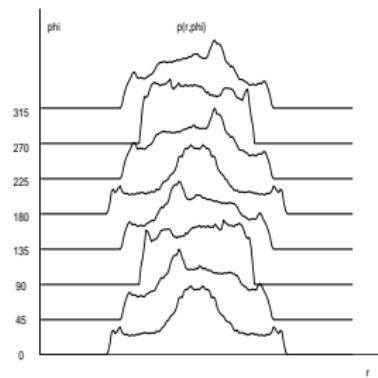
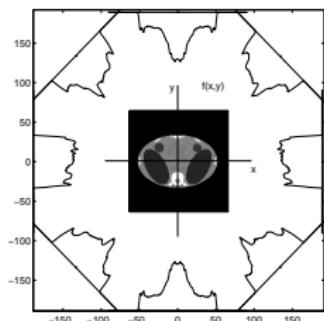
$$g(s) = \int f(r) \exp\{-js.r\} dr$$

$$h(s, r) = \exp\{-js.r\} \longrightarrow w(s, r) = \exp\{+js.r\}$$

$$\hat{f}(r) = \int g(s) \exp\{+js.r\} ds$$

- Known classical solutions for specific expressions of  $h(s, r)$ :
  - 1D cases: 1D Fourier, Hilbert, Weil, Melin, ...
  - 2D cases: 2D Fourier, Radon, ...

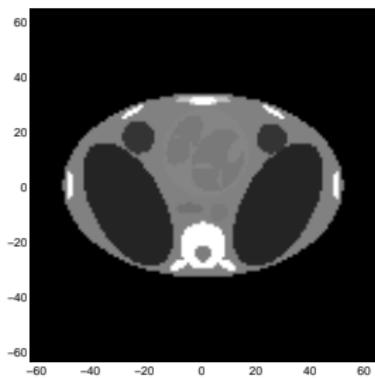
# X ray Tomography



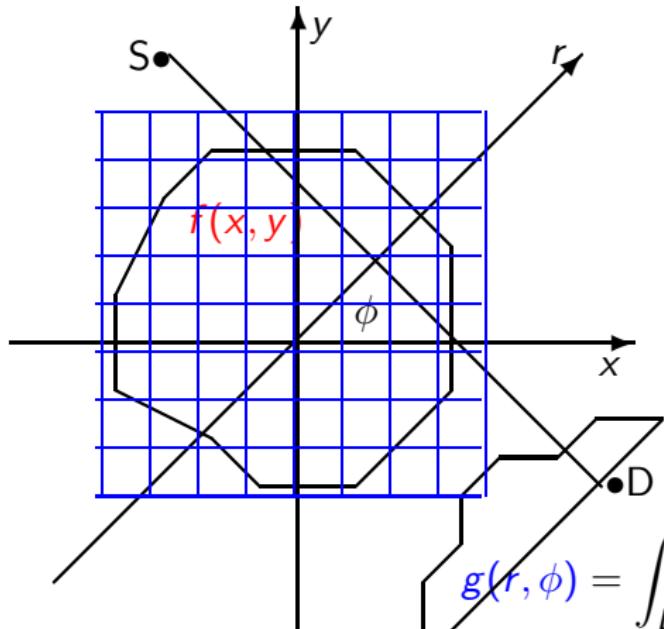
$$g(r, \phi) = -\ln \left( \frac{I}{I_0} \right) = \int_{L_{r,\phi}} f(x, y) \, dx$$
$$g(r, \phi) = \iint_D f(x, y) \delta(r - x \cos \phi - y \sin \phi) \, dx \, dy$$



IRT  
?  
====>



# Analytical Inversion methods



Radon:

$$g(r, \phi) = \iint_D f(x, y) \delta(r - x \cos \phi - y \sin \phi) \, dx \, dy$$

$$f(x, y) = \left( -\frac{1}{2\pi^2} \right) \int_0^\pi \int_{-\infty}^{+\infty} \frac{\frac{\partial}{\partial r} g(r, \phi)}{(r - x \cos \phi - y \sin \phi)} \, dr \, d\phi$$

# Filtered Backprojection method

$$f(x, y) = \left( -\frac{1}{2\pi^2} \right) \int_0^\pi \int_{-\infty}^{+\infty} \frac{\frac{\partial}{\partial r} g(r, \phi)}{(r - x \cos \phi - y \sin \phi)} dr d\phi$$

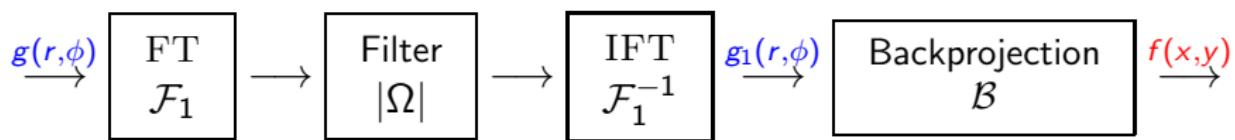
Derivation  $\mathcal{D}$  :  $\bar{g}(r, \phi) = \frac{\partial g(r, \phi)}{\partial r}$

Hilbert Transform  $\mathcal{H}$  :  $g_1(r', \phi) = \frac{1}{\pi} \int_0^\infty \frac{\bar{g}(r, \phi)}{(r - r')} dr$

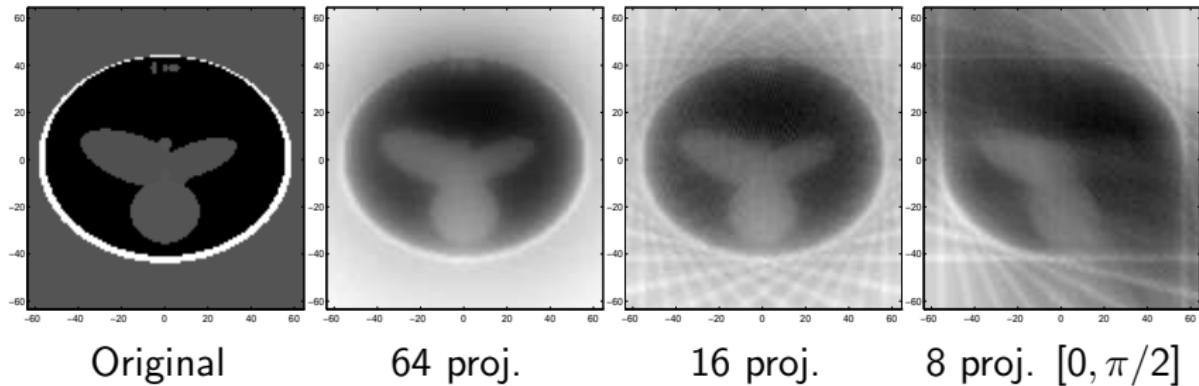
Backprojection  $\mathcal{B}$  :  $f(x, y) = \frac{1}{2\pi} \int_0^\pi g_1(r' = x \cos \phi + y \sin \phi, \phi) d\phi$

$$f(x, y) = \mathcal{B} \mathcal{H} \mathcal{D} g(r, \phi) = \mathcal{B} \mathcal{F}_1^{-1} |\Omega| \mathcal{F}_1 g(r, \phi)$$

- Backprojection of filtered projections:

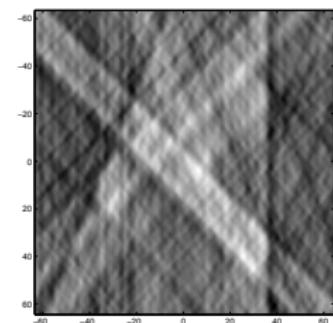
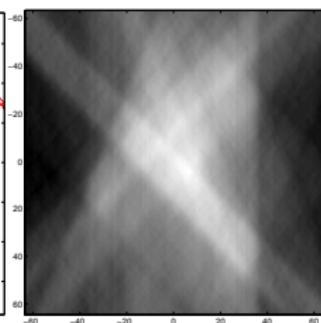
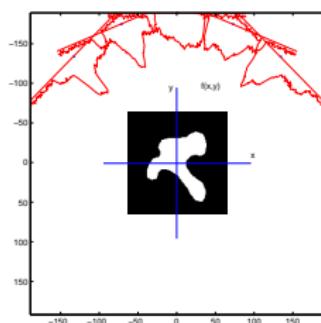
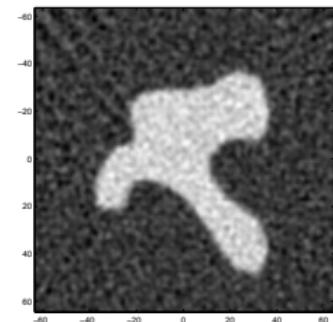
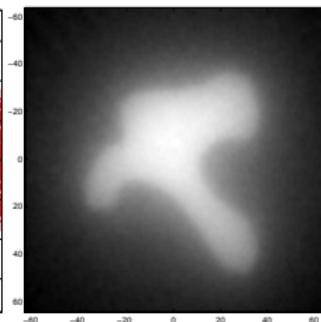
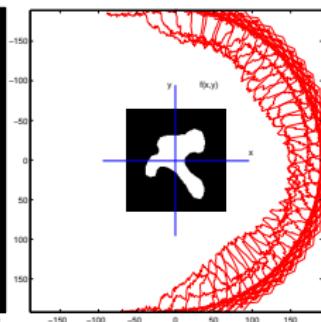
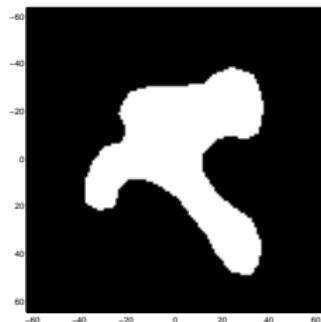


## Limitations : Limited angle or noisy data



- ▶ Limited angle or noisy data
- ▶ Accounting for detector size
- ▶ Other measurement geometries: fan beam, ...

## Limitations : Limited angle or noisy data



Original

Data

Backprojection

Filtered Backprojection

## Parametric methods

- ▶  $f(\mathbf{r})$  is described in a parametric form with a very few number of parameters  $\boldsymbol{\theta}$  and one searches  $\hat{\boldsymbol{\theta}}$  which minimizes a criterion such as:
- ▶ Least Squares (LS): 
$$Q(\boldsymbol{\theta}) = \sum_i |g_i - [\mathcal{H} f(\boldsymbol{\theta})]_i|^2$$
- ▶ Robust criteria :  
with different functions  $\phi$  
$$Q(\boldsymbol{\theta}) = \sum_i \phi(|g_i - [\mathcal{H} f(\boldsymbol{\theta})]_i|)$$
( $L_1$ , Hubert, ...).
- ▶ Likelihood : 
$$\mathcal{L}(\boldsymbol{\theta}) = -\ln p(g|\boldsymbol{\theta})$$
- ▶ Penalized likelihood : 
$$\mathcal{L}(\boldsymbol{\theta}) = -\ln p(g|\boldsymbol{\theta}) + \lambda \Omega(\boldsymbol{\theta})$$

Examples:

- ▶ Spectrometry:  $f(t)$  modelled as a sum of gaussians  
$$f(t) = \sum_{k=1}^K a_k \mathcal{N}(t|\mu_k, \nu_k) \quad \boldsymbol{\theta} = \{a_k, \mu_k, \nu_k\}$$
- ▶ Tomography in CND:  $f(x, y)$  is modelled as a superposition of circular or elliptical discs 
$$\boldsymbol{\theta} = \{a_k, \mu_k, r_k\}$$

## Non parametric methods

$$g(s_i) = \int h(s_i, r) f(r) dr + \epsilon(s_i), \quad i = 1, \dots, M$$

- $f(r)$  is assumed to be well approximated by

$$f(r) \simeq \sum_{j=1}^N f_j b_j(r)$$

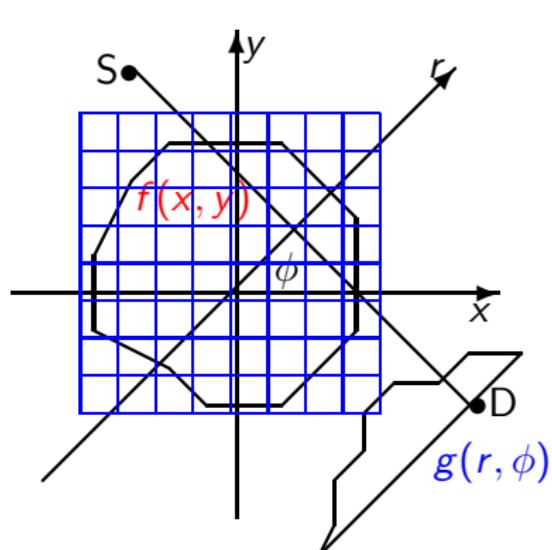
with  $\{b_j(r)\}$  a basis or any other set of known functions

$$g(s_i) = g_i \simeq \sum_{j=1}^N f_j \int h(s_i, r) b_j(r) dr, \quad i = 1, \dots, M$$

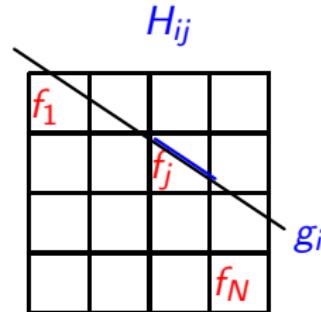
$$\mathbf{g} = \mathbf{H} \mathbf{f} + \boldsymbol{\epsilon} \quad \text{with} \quad H_{ij} = \int h(s_i, r) b_j(r) dr$$

- $H$  is huge dimensional
- LS solution :  $\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{Q(\mathbf{f})\}$  with  
$$Q(\mathbf{f}) = \sum_i |g_i - [\mathbf{H} \mathbf{f}]_i|^2 = \|\mathbf{g} - \mathbf{H} \mathbf{f}\|^2$$
  
does not give satisfactory result.

# Algebraic methods: Discretization



$$g(r, \phi) = \int_L f(x, y) \, dl$$



$$f(x, y) = \sum_j f_j b_j(x, y)$$

$$b_j(x, y) = \begin{cases} 1 & \text{if } (x, y) \in \text{ pixel } j \\ 0 & \text{else} \end{cases}$$

$$g_i = \sum_{j=1}^N H_{ij} f_j + \epsilon_i$$

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon}$$

# Inversion: Deterministic methods

## Data matching

- ▶ Observation model

$$g_i = h_i(\mathbf{f}) + \epsilon_i, \quad i = 1, \dots, M \longrightarrow \mathbf{g} = \mathbf{H}(\mathbf{f}) + \boldsymbol{\epsilon}$$

- ▶ Misatch between data and output of the model  $\Delta(\mathbf{g}, \mathbf{H}(\mathbf{f}))$

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{\Delta(\mathbf{g}, \mathbf{H}(\mathbf{f}))\}$$

- ▶ Examples:

– LS       $\Delta(\mathbf{g}, \mathbf{H}(\mathbf{f})) = \|\mathbf{g} - \mathbf{H}(\mathbf{f})\|^2 = \sum_i |g_i - h_i(\mathbf{f})|^2$

–  $L_p$        $\Delta(\mathbf{g}, \mathbf{H}(\mathbf{f})) = \|\mathbf{g} - \mathbf{H}(\mathbf{f})\|^p = \sum_i |g_i - h_i(\mathbf{f})|^p, \quad 1 < p < 2$

– KL       $\Delta(\mathbf{g}, \mathbf{H}(\mathbf{f})) = \sum_i g_i \ln \frac{g_i}{h_i(\mathbf{f})}$

- ▶ In general, does not give satisfactory results for inverse problems.

# Regularization theory

Inverse problems = III posed problems  
→ Need for prior information

Functional space (Tikhonov):

$$\mathbf{g} = \mathcal{H}(\mathbf{f}) + \epsilon \rightarrow J(\mathbf{f}) = \|\mathbf{g} - \mathcal{H}(\mathbf{f})\|_2^2 + \lambda \|\mathcal{D}\mathbf{f}\|_2^2$$

Finite dimensional space (Philips & Towlmey):  $\mathbf{g} = \mathbf{H}(\mathbf{f}) + \epsilon$

- Minimum norm LS (MNLS):  $J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}(\mathbf{f})\|^2 + \lambda \|\mathbf{f}\|^2$
- Classical regularization:  $J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}(\mathbf{f})\|^2 + \lambda \|\mathcal{D}\mathbf{f}\|^2$
- More general regularization:

$$J(\mathbf{f}) = \mathcal{Q}(\mathbf{g} - \mathbf{H}(\mathbf{f})) + \lambda \Omega(\mathcal{D}\mathbf{f})$$

or

$$J(\mathbf{f}) = \Delta_1(\mathbf{g}, \mathbf{H}(\mathbf{f})) + \lambda \Delta_2(\mathbf{f}, \mathbf{f}_\infty)$$

**Limitations:**

- Errors are considered implicitly white and Gaussian
- Limited prior information on the solution
- Lack of tools for the determination of the hyperparameters

# Inversion: Probabilistic methods

Taking account of errors and uncertainties → Probability theory

- ▶ Maximum Likelihood (ML)
- ▶ Minimum Inaccuracy (MI)
- ▶ Probability Distribution Matching (PDM)
- ▶ Maximum Entropy (ME) and Information Theory (IT)
- ▶ Bayesian Inference (BAYES)

## Advantages:

- ▶ Explicit account of the errors and noise
- ▶ A large class of priors via explicit or implicit modeling
- ▶ A coherent approach to combine information content of the data and priors

## Limitations:

- ▶ Practical implementation and cost of calculation

# Bayesian estimation approach

$$\mathcal{M} : \quad \mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon}$$

- ▶ Observation model  $\mathcal{M}$  + Hypothesis on the noise  $\boldsymbol{\epsilon}$   $\rightarrow p(\mathbf{g}|\mathbf{f}; \mathcal{M}) = p_{\boldsymbol{\epsilon}}(\mathbf{g} - \mathbf{H}\mathbf{f})$
- ▶ A priori information  $p(\mathbf{f}|\mathcal{M})$
- ▶ Bayes : 
$$p(\mathbf{f}|\mathbf{g}; \mathcal{M}) = \frac{p(\mathbf{g}|\mathbf{f}; \mathcal{M}) p(\mathbf{f}|\mathcal{M})}{p(\mathbf{g}|\mathcal{M})}$$

## Link with regularization :

Maximum A Posteriori (MAP) :

$$\begin{aligned}\hat{\mathbf{f}} &= \arg \max_{\mathbf{f}} \{p(\mathbf{f}|\mathbf{g})\} = \arg \max_{\mathbf{f}} \{p(\mathbf{g}|\mathbf{f}) p(\mathbf{f})\} \\ &= \arg \min_{\mathbf{f}} \{-\ln p(\mathbf{g}|\mathbf{f}) - \ln p(\mathbf{f})\}\end{aligned}$$

with  $Q(\mathbf{g}, \mathbf{H}\mathbf{f}) = -\ln p(\mathbf{g}|\mathbf{f})$  and  $\lambda\Omega(\mathbf{f}) = -\ln p(\mathbf{f})$

## Case of linear models and Gaussian priors

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon}$$

- ▶ Hypothesis on the noise:  $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_\epsilon^2 \mathbf{I}) \longrightarrow p(\mathbf{g}|\mathbf{f}) \propto \exp\left\{-\frac{1}{2\sigma_\epsilon^2} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2\right\}$
- ▶ Hypothesis on  $\mathbf{f}$ :  $\mathbf{f} \sim \mathcal{N}(0, \sigma_f^2 (\mathbf{D}^t \mathbf{D})^{-1}) \longrightarrow p(\mathbf{f}) \propto \exp\left\{-\frac{1}{2\sigma_f^2} \|\mathbf{D}\mathbf{f}\|^2\right\}$
- ▶ A posteriori:  
$$p(\mathbf{f}|\mathbf{g}) \propto \exp\left\{-\frac{1}{2\sigma_\epsilon^2} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 - \frac{1}{2\sigma_f^2} \|\mathbf{D}\mathbf{f}\|^2\right\}$$
- ▶ MAP :  $\widehat{\mathbf{f}} = \arg \max_{\mathbf{f}} \{p(\mathbf{f}|\mathbf{g})\} = \arg \min_{\mathbf{f}} \{J(\mathbf{f})\}$   
with  $J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \lambda \|\mathbf{D}\mathbf{f}\|^2, \quad \lambda = \frac{\sigma_\epsilon^2}{\sigma_f^2}$
- ▶ Advantage : characterization of the solution

$$\mathbf{f}|\mathbf{g} \sim \mathcal{N}(\widehat{\mathbf{f}}, \widehat{\mathbf{P}}) \text{ with } \widehat{\mathbf{f}} = \widehat{\mathbf{P}} \mathbf{H}^t \mathbf{g}, \quad \widehat{\mathbf{P}} = (\mathbf{H}^t \mathbf{H} + \lambda \mathbf{D}^t \mathbf{D})^{-1}$$

## MAP estimation with other priors:

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{J(\mathbf{f})\} \quad \text{with} \quad J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \lambda \Omega(\mathbf{f})$$

### Separable priors:

- ▶ Gaussian:  $p(f_j) \propto \exp \{-\alpha|f_j|^2\} \longrightarrow \Omega(\mathbf{f}) = \alpha \sum_j |f_j|^2$
- ▶ Gamma:  $p(f_j) \propto f_j^\alpha \exp \{-\beta f_j\} \longrightarrow \Omega(\mathbf{f}) = \alpha \sum_j \ln f_j + \beta f_j$
- ▶ Beta:  
 $p(f_j) \propto f_j^\alpha (1-f_j)^\beta \longrightarrow \Omega(\mathbf{f}) = \alpha \sum_j \ln f_j + \beta \sum_j \ln(1-f_j)$
- ▶ Generalized Gaussian:  
 $p(f_j) \propto \exp \{-\alpha|f_j|^p\}, \quad 1 < p < 2 \longrightarrow \Omega(\mathbf{f}) = \alpha \sum_j |f_j|^p,$

### Markovian models:

$$p(f_j|\mathbf{f}) \propto \exp \left\{ -\alpha \sum_{i \in N_j} \phi(f_j, f_i) \right\} \rightarrow \Omega(\mathbf{f}) = \alpha \sum_j \sum_{i \in N_j} \phi(f_j, f_i),$$

## MAP estimation with markovien priors:

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{J(\mathbf{f})\} \quad \text{with} \quad J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \lambda \Omega(\mathbf{f})$$

$$\Omega(\mathbf{f}) = \sum_j \phi(f_j - f_{j-1})$$

with  $\phi(t)$  :

Convex functions:

$$|t|^\alpha, \sqrt{1+t^2} - 1, \log(\cosh(t)), \quad \begin{cases} t^2 & |t| \leq T \\ 2T|t| - T^2 & |t| > T \end{cases}$$

or Non convex functions:

$$\log(1+t^2), \quad \frac{t^2}{1+t^2}, \quad \arctan(t^2), \quad \begin{cases} t^2 & |t| \leq T \\ T^2 & |t| > T \end{cases}$$

## Main advantages of the Bayesian approach

- ▶ MAP = Regularization
- ▶ Posterior mean ? Marginal MAP ?
- ▶ More information in the posterior law than only its mode or its mean
- ▶ Meaning and tools for estimating hyper parameters
- ▶ Meaning and tools for model selection
- ▶ More specific and specialized priors, particularly through the hidden variables
- ▶ More computational tools:
  - ▶ Expectation-Maximization for computing the maximum likelihood parameters
  - ▶ MCMC for posterior exploration
  - ▶ Variational Bayes for analytical computation of the posterior marginals
  - ▶ ...

## Full Bayesian approach

$$\mathcal{M} : \quad \mathbf{g} = \mathbf{H} \mathbf{f} + \boldsymbol{\epsilon}$$

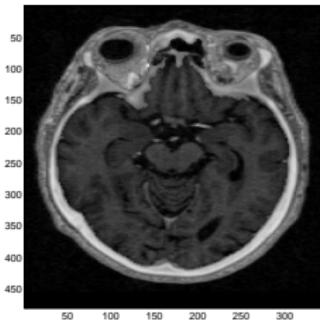
- ▶ Forward & errors model:  $\rightarrow p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}_1; \mathcal{M})$
- ▶ Prior models  $\rightarrow p(\mathbf{f}|\boldsymbol{\theta}_2; \mathcal{M})$
- ▶ Hyperparameters  $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \rightarrow p(\boldsymbol{\theta}|\mathcal{M})$
- ▶ Bayes:  $\rightarrow p(\mathbf{f}, \boldsymbol{\theta}|\mathbf{g}; \mathcal{M}) = \frac{p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}; \mathcal{M}) p(\mathbf{f}|\boldsymbol{\theta}; \mathcal{M}) p(\boldsymbol{\theta}|\mathcal{M})}{p(\mathbf{g}|\mathcal{M})}$
- ▶ Joint MAP:  $(\hat{\mathbf{f}}, \hat{\boldsymbol{\theta}}) = \arg \max_{(\mathbf{f}, \boldsymbol{\theta})} \{p(\mathbf{f}, \boldsymbol{\theta}|\mathbf{g}; \mathcal{M})\}$
- ▶ Marginalization: 
$$\begin{cases} p(\mathbf{f}|\mathbf{g}; \mathcal{M}) &= \int p(\mathbf{f}, \boldsymbol{\theta}|\mathbf{g}; \mathcal{M}) d\boldsymbol{\theta} \\ p(\boldsymbol{\theta}|\mathbf{g}; \mathcal{M}) &= \int p(\mathbf{f}, \boldsymbol{\theta}|\mathbf{g}; \mathcal{M}) d\mathbf{f} \end{cases}$$
- ▶ Posterior means: 
$$\begin{cases} \hat{\mathbf{f}} &= \int \mathbf{f} p(\mathbf{f}, \boldsymbol{\theta}|\mathbf{g}; \mathcal{M}) d\boldsymbol{\theta} d\mathbf{f} \\ \hat{\boldsymbol{\theta}} &= \int \boldsymbol{\theta} p(\mathbf{f}, \boldsymbol{\theta}|\mathbf{g}; \mathcal{M}) d\mathbf{f} d\boldsymbol{\theta} \end{cases}$$
- ▶ Evidence of the model:

$$p(\mathbf{g}|\mathcal{M}) = \iint p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}; \mathcal{M}) p(\mathbf{f}|\boldsymbol{\theta}; \mathcal{M}) p(\boldsymbol{\theta}|\mathcal{M}) d\mathbf{f} d\boldsymbol{\theta}$$

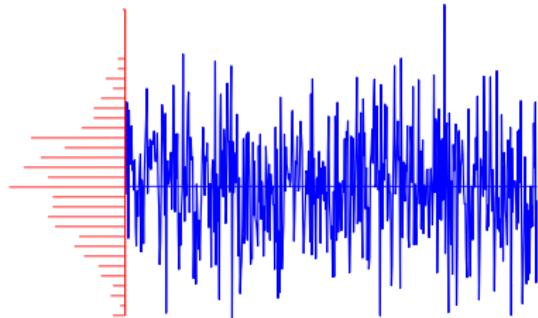
# Two main steps in the Bayesian approach

- ▶ Prior modeling
  - ▶ Separable:  
Gaussian, Generalized Gaussian, Gamma,  
mixture of Gaussians, mixture of Gammas, ...
  - ▶ Markovian: Gauss-Markov, GGM, ...
  - ▶ Separable or Markovian with **hidden variables**  
(contours, region labels)
- ▶ Choice of the estimator and computational aspects
  - ▶ MAP, Posterior mean, Marginal MAP
  - ▶ MAP needs **optimization** algorithms
  - ▶ Posterior mean needs **integration** methods
  - ▶ Marginal MAP needs integration and optimization
  - ▶ Approximations:
    - ▶ Gaussian approximation (Laplace)
    - ▶ Numerical exploration MCMC
    - ▶ Variational Bayes (Separable approximation)

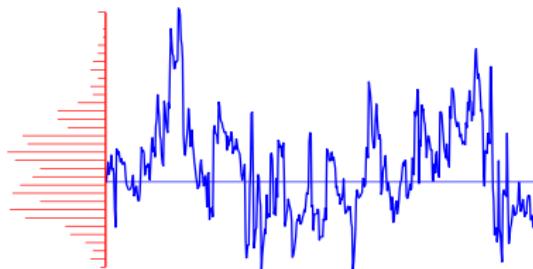
# Which images I am looking for?



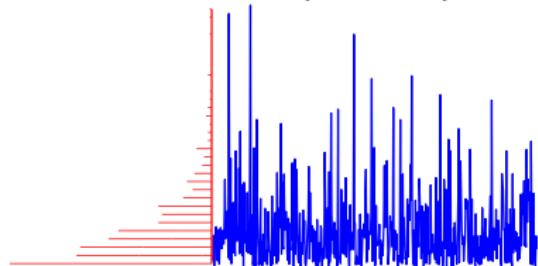
# Which signals I am looking for?



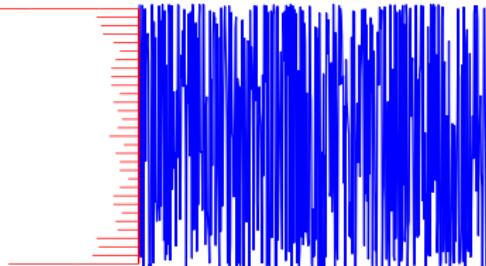
Gaussian

$$p(f_j) \propto \exp \{-\alpha |f_j|^2\}$$


Generalized Gaussian

$$p(f_j) \propto \exp \{-\alpha |f_j|^p\}, \quad 1 \leq p \leq 2$$


Gamma

$$p(f_j) \propto f_j^\alpha \exp \{-\beta f_j\}$$


Beta

$$p(f_j) \propto f_j^\alpha (1 - f_j)^\beta$$

## Different prior models for signals and images

- Separable  $p(\mathbf{f}) = \prod_j p_j(f_j) \propto \exp \left\{ -\beta \sum_j \phi(f_j) \right\}$

$$p(\mathbf{f}) \propto \exp \left\{ -\beta \sum_{\mathbf{r} \in \mathcal{R}} \phi(f(\mathbf{r})) \right\}$$

- Markoviens (simple)  $p(f_j | f_{j-1}) \propto \exp \{ -\beta \phi(f_j - f_{j-1}) \}$

$$p(\mathbf{f}) \propto \exp \left\{ -\beta \sum_{\mathbf{r} \in \mathcal{R}} \sum_{\mathbf{r}' \in \mathcal{V}(\mathbf{r})} \phi(f(\mathbf{r}), f(\mathbf{r}')) \right\}$$

- Markovien with hidden variables  
 $z(\mathbf{r})$  (lines, contours, regions)

$$p(\mathbf{f} | \mathbf{z}) \propto \exp \left\{ -\beta \sum_{\mathbf{r} \in \mathcal{R}} \sum_{\mathbf{r}' \in \mathcal{V}(\mathbf{r})} \phi(f(\mathbf{r}), f(\mathbf{r}'), z(\mathbf{r}), z(\mathbf{r}')) \right\}$$

## Different prior models for images: Separable

- Gaussian:

$$p(f_j) \propto \exp \{-\alpha |f_j|^2\} \rightarrow \Omega(\mathbf{f}) = \alpha \sum_j |f_j|^2$$

- Generalized Gaussian (GG):

$$p(f_j) \propto \exp \{-\alpha |f_j|^p\}, \quad 1 \leq p \leq 2 \rightarrow \Phi(\mathbf{f}) = \alpha \sum_j |f_j|^p,$$

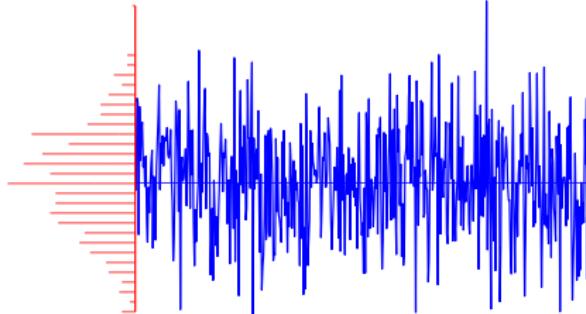
- Gamma:  $f_j > 0$

$$p(f_j) \propto f_j^\alpha \exp \{-\beta f_j\} \rightarrow \Omega(\mathbf{f}) = \alpha \sum_j \ln f_j + \beta \sum_j f_j,$$

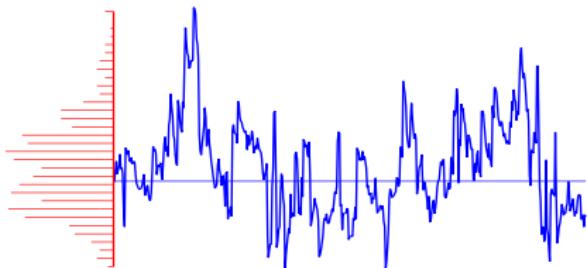
- Beta:  $0 < f_j < 1$

$$p(f_j) \propto f_j^\alpha (1 - f_j)^\beta \rightarrow \Omega(\mathbf{f}) = \alpha \sum_j \ln f_j + \beta \sum_j \ln(1 - f_j),$$

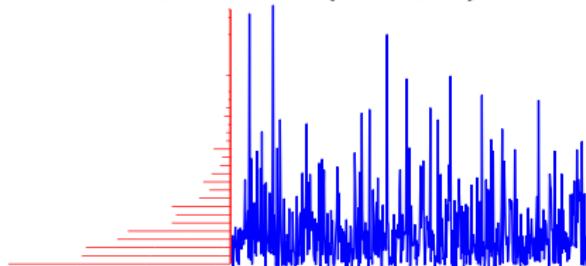
## Different prior models for images: Separable



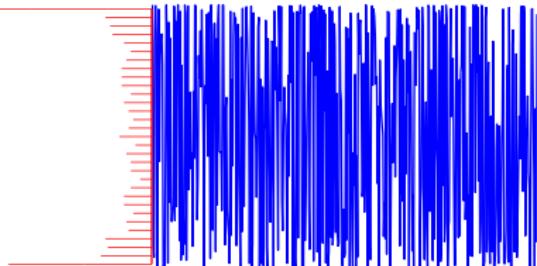
Gaussian  
 $p(f_j) \propto \exp \{-\alpha |f_j|^2\}$



Generalized Gaussian  
 $p(f_j) \propto \exp \{-\alpha |f_j|^p\}, \quad 1 \leq p \leq 2$



Gamma  
 $p(f_j) \propto f_j^\alpha \exp \{-\beta f_j\}$



Beta  
 $p(f_j) \propto f_j^\alpha (1 - f_j)^\beta$

## Different prior models: Simple Markovian

$$p(f_j | \mathbf{f}) \propto \exp \left\{ -\alpha \sum_{i \in V_j} \phi(f_j, f_i) \right\} \rightarrow \Phi(\mathbf{f}) = \alpha \sum_j \sum_{i \in V_j} \phi(f_j, f_i)$$

- 1D case and one neighbor  $V_j = j - 1$ :

$$\Phi(\mathbf{f}) = \alpha \sum_j \phi(f_j - f_{j-1})$$

- 1D Case and two neighbors  $V_j = \{j - 1, j + 1\}$ :

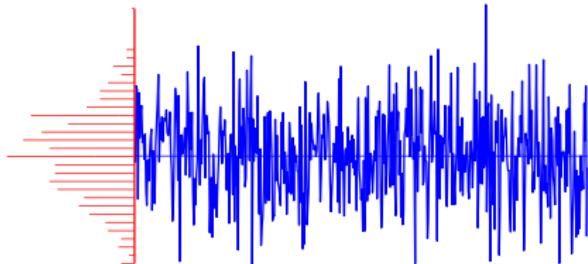
$$\Phi(\mathbf{f}) = \alpha \sum_j \phi(f_j - \beta(f_{j-1} + f_{j+1}))$$

- 2D case with 4 neighbors:

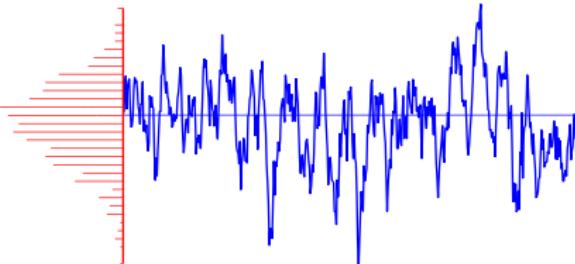
$$\Phi(\mathbf{f}) = \alpha \sum_{\mathbf{r} \in \mathcal{R}} \phi \left( f(\mathbf{r}) - \beta \sum_{\mathbf{r}' \in \mathcal{V}(\mathbf{r})} f(\mathbf{r}') \right)$$

- $\phi(t) = |t|^\gamma$ : Generalized Gaussian

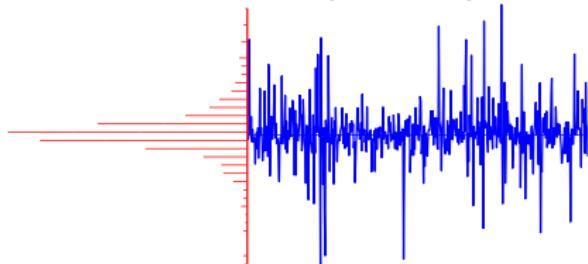
## Different prior models: Simple Markovian



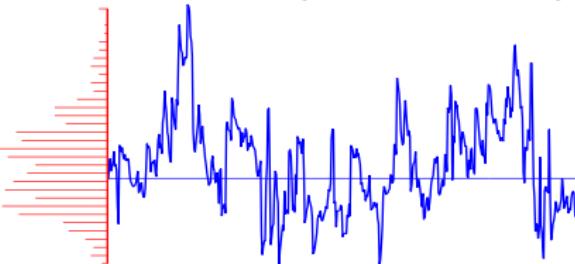
IID Gaussian  
 $p(f_j) \propto \exp \{-\alpha |f_j|^2\}$



Gauss-Markov  
 $p(f_j|f_{j-1}) \propto \exp \{-\alpha |f_j - f_{j-1}|^2\}$

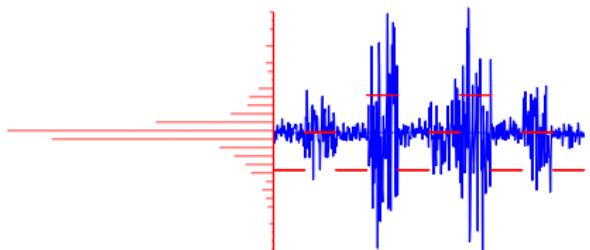


IID GG  
 $p(f_j) \propto \exp \{-\alpha |f_j|^p\}$

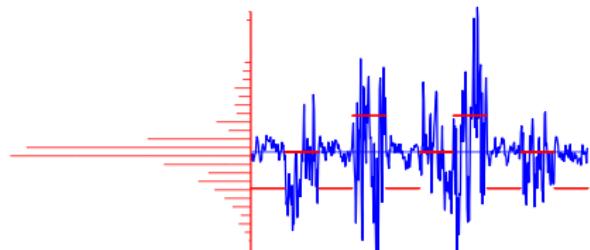


Markovian GG  
 $p(f_j|f_{j-1}) \propto \exp \{-\alpha |f_j - f_{j-1}|^p\}$

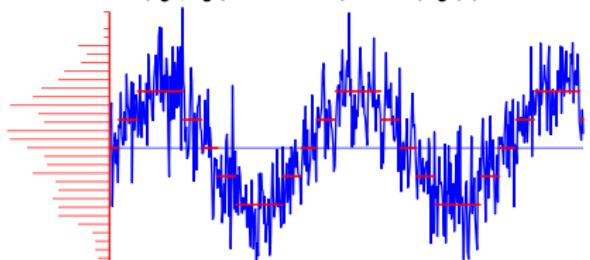
## Different prior models: Non-stationnary signals



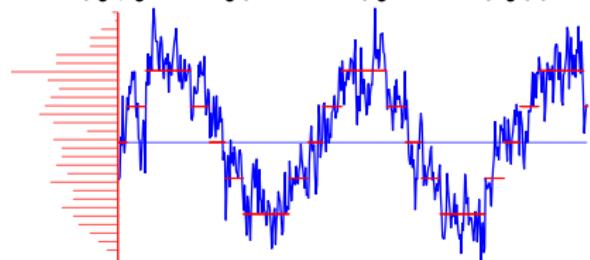
Modulated Variances IID  
 $p(f_j|z_j) = \mathcal{N}(0, v(z_j))$



Modulated Variances Gauss-Markov  
 $p(f_j|f_{j-1}, z_j) = \mathcal{N}(f_{j-1}, v(z_j))$

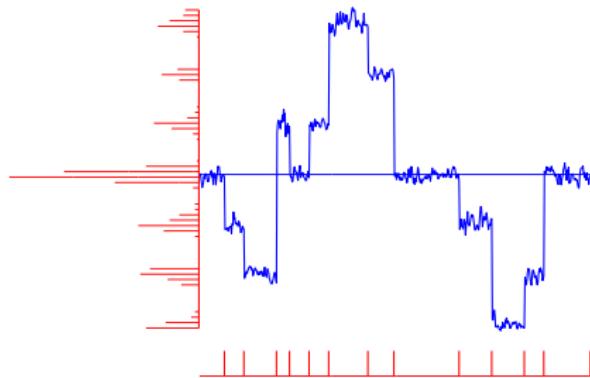


Modulated amplituds IID  
 $p(f_j|z_j) = \mathcal{N}(a(z_j), 1)$



Modulated amplituds Gauss-Markov  
 $p(f_j|f_{j-1}, z_j) = \mathcal{N}(a(f_{j-1}, z_j), 1)$

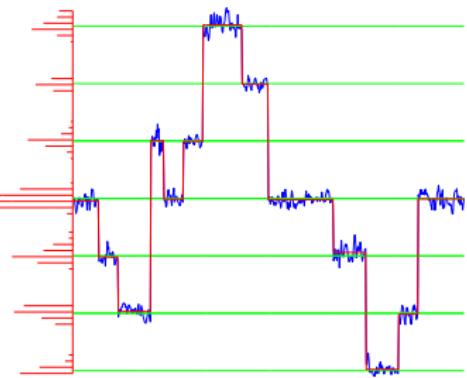
# Different prior models: Markovian with hidden variables



Piecewise Gaussians

(contours hidden variables)

$$p(f_j|q_j, f_{j-1}) = \mathcal{N}((1 - q_j)f_{j-1}, \sigma_f^2)$$



Mixture of Gaussians (MoG)

(regions labels hidden variables)

$$p(f_j|z_j = k) = \mathcal{N}(m_k, \sigma_k^2) \text{ & } z_j \text{ markovian}$$

$$p(\mathbf{f}|\mathbf{q}) \propto \exp \left\{ -\alpha \sum_j |f_j - (1 - q_j)f_{j-1}|^2 \right\}$$

$$p(\mathbf{f}|\mathbf{z}) \propto \exp \left\{ -\alpha \sum_k \sum_{j \in \mathcal{R}_k} \left( \frac{f_j - m_k}{\sigma_k} \right)^2 \right\}$$

## Particular case of Gauss-Markov models

$$\begin{cases} \mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon} \text{ with} \\ \mathbf{f} \sim \mathcal{N}(0, \sigma_f^2(\mathbf{D}^t\mathbf{D})^{-1}) \end{cases} = \begin{cases} \mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon} \\ \mathbf{f} = \mathbf{C}\mathbf{f} + \mathbf{z} \text{ with } \mathbf{z} \sim \mathcal{N}(0, \sigma_f^2 \mathbf{I}) \\ \text{and } \mathbf{D} = (\mathbf{I} - \mathbf{C}) \end{cases}$$

$$\mathbf{f}|\mathbf{g} \sim \mathcal{N}(\hat{\mathbf{f}}, \hat{\mathbf{P}}) \text{ with } \hat{\mathbf{f}} = \hat{\mathbf{P}}\mathbf{H}^t\mathbf{g}, \quad \hat{\mathbf{P}} = (\mathbf{H}^t\mathbf{H} + \lambda\mathbf{D}^t\mathbf{D})^{-1}$$

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{ J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \lambda \|\mathbf{D}\mathbf{f}\|^2 \}$$

$$\begin{cases} \mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon} \\ \text{with } \mathbf{f} \sim \mathcal{N}(0, \sigma_f^2(\mathbf{D}\mathbf{D}^t)) \end{cases} = \begin{cases} \mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon} \\ \mathbf{f} = \mathbf{D}\mathbf{z} \text{ with } \mathbf{z} \sim \mathcal{N}(0, \sigma_f^2 \mathbf{I}) \end{cases}$$

$$\mathbf{z}|\mathbf{g} \sim \mathcal{N}(\hat{\mathbf{z}}, \hat{\mathbf{P}}) \text{ with } \hat{\mathbf{z}} = \hat{\mathbf{P}}\mathbf{D}^t\mathbf{H}^t\mathbf{g}, \quad \hat{\mathbf{P}} = (\mathbf{D}^t\mathbf{H}^t\mathbf{H}\mathbf{D} + \lambda\mathbf{I})^{-1}$$

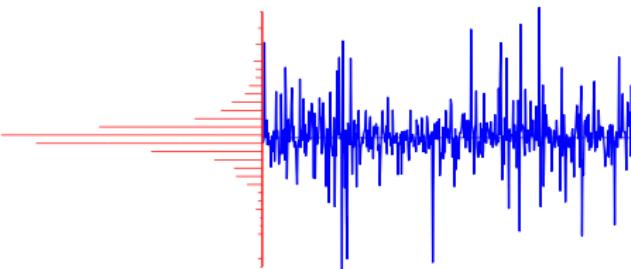
$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z}} \{ J(\mathbf{z}) = \|\mathbf{g} - \mathbf{H}\mathbf{D}\mathbf{z}\|^2 + \lambda \|\mathbf{z}\|^2 \} \longrightarrow \hat{\mathbf{f}} = \mathbf{D}\hat{\mathbf{z}}$$

$\mathbf{z}$  Decomposition coeff on a basis (column of  $\mathbf{D}$ )

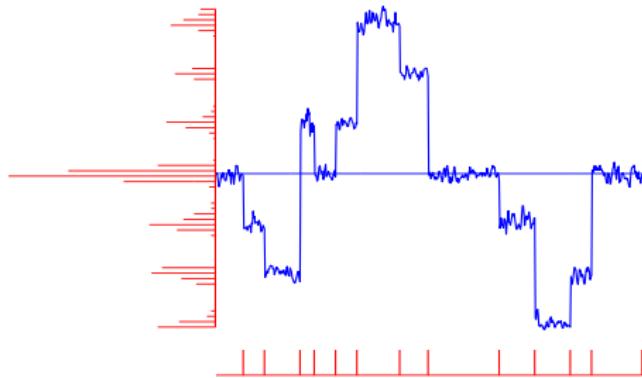
# Which images I am looking for?



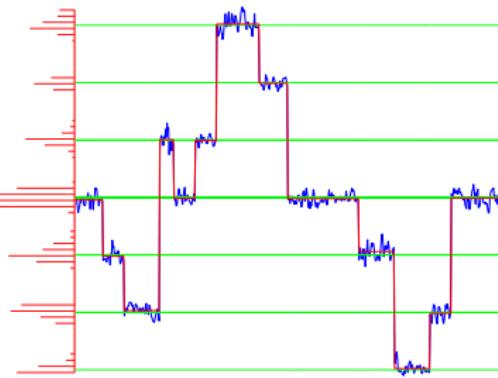
Gauss-Markov



Generalized GM



Piecewise Gaussian



Mixture of GM

# Markovien prior models for images

$$\Omega(\mathbf{f}) = \sum_j \phi(f_j - f_{j-1})$$

- ▶ Gauss-Markov :  $\phi(t) = |t|^2$
- ▶ Generalized Gauss-Markov :  $\phi(t) = |t|^\alpha$
- ▶ Picewise Gauss-Markov or GGM :  $\phi(t) = \begin{cases} t^2 & |t| \leq T \\ T^2 & |t| > T \end{cases}$   
or equivalently :

$$\Omega(\mathbf{f}|\mathbf{q}) = \sum_j (1 - q_j) \phi(f_j - f_{j-1})$$

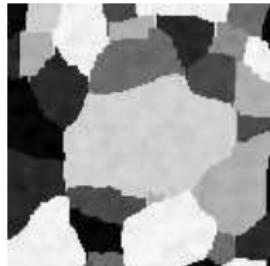
$\mathbf{q}$  line process (contours)

- ▶ Mixture of Gaussians :

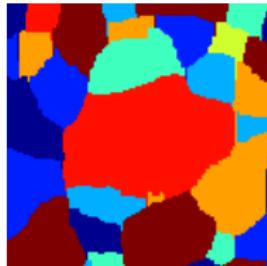
$$\Omega(\mathbf{f}|\mathbf{z}) = \sum_k \sum_{\{j:z_j=k\}} \left( \frac{f_j - m_k}{v_k} \right)^2$$

$\mathbf{z}$  region labels process.

# Gauss-Markov-Potts prior models for images



$f(r)$



$z(r)$



$$c(r) = 1 - \delta(z(r) - z(r'))$$

$$p(f(r)|z(r) = k, m_k, v_k) = \mathcal{N}(m_k, v_k)$$

$$p(f(r)) = \sum_k P(z(r) = k) \mathcal{N}(m_k, v_k) \quad \text{Mixture of Gaussians}$$

- ▶ Separable iid hidden variables:  $p(z) = \prod_r p(z(r))$
- ▶ Markovian hidden variables:  $p(z)$  Potts-Markov:

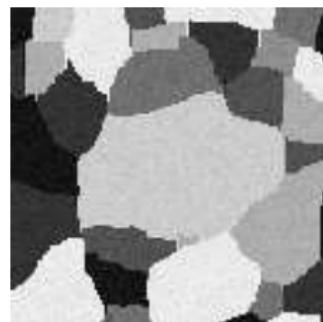
$$p(z(r)|z(r'), r' \in \mathcal{V}(r)) \propto \exp \left\{ \gamma \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}$$
$$p(z) \propto \exp \left\{ \gamma \sum_{r \in \mathcal{R}} \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}$$

## Four different cases

To each pixel of the image is associated 2 variables  $f(r)$  and  $z(r)$

- ▶  $f|z$  Gaussian iid,  $z$  iid :

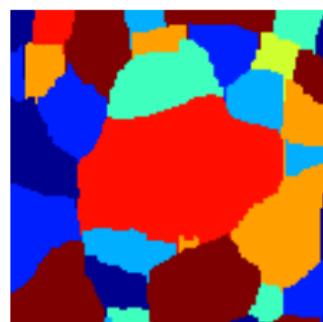
Mixture of Gaussians



$f(r)$

- ▶  $f|z$  Gauss-Markov,  $z$  iid :

Mixture of Gauss-Markov

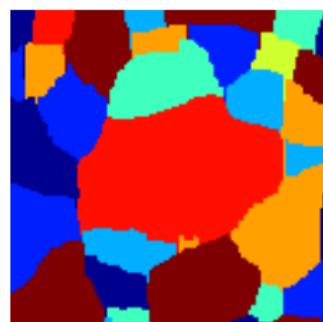


$z(r)$

- ▶  $f|z$  Gaussian iid,  $z$  Potts-Markov :

Mixture of Independent Gaussians

(MIG with Hidden Potts)



$z(r)$

- ▶  $f|z$  Markov,  $z$  Potts-Markov :

Mixture of Gauss-Markov

(MGM with hidden Potts)

Case 1:  $f|z$  Gaussian iid,  $z$  iid

Independent Mixture of Independent Gaussians (IMIG):

$$p(f(r)|z(r) = k) = \mathcal{N}(m_k, v_k), \quad \forall r \in \mathcal{R}$$

$$p(f(r)) = \sum_{k=1}^K \alpha_k \mathcal{N}(m_k, v_k), \text{ with } \sum_k \alpha_k = 1.$$

$$p(z) = \prod_r p(z(r) = k) = \prod_r \alpha_k = \prod_k \alpha_k^{n_k}$$

Noting

$$m_z(r) = m_k, v_z(r) = v_k, \alpha_z(r) = \alpha_k, \forall r \in \mathcal{R}_k$$

we have:

$$p(f|z) = \prod_{r \in \mathcal{R}} \mathcal{N}(m_z(r), v_z(r))$$

$$p(z) = \prod_r \alpha_z(r) = \prod_k \alpha_k^{\sum_{r \in \mathcal{R}} \delta(z(r) - k)} = \prod_k \alpha_k^{n_k}$$

Case 2:  $f|z$  Gauss-Markov,  $z$  iid

Independent Mixture of Gauss-Markov (IMGM):

$$p(f(r)|z(r), z(r'), f(r'), r' \in \mathcal{V}(r)) = \mathcal{N}(\mu_z(r), v_z(r)), \forall r \in \mathcal{R}$$

$$\begin{aligned}\mu_z(r) &= \frac{1}{|\mathcal{V}(r)|} \sum_{r' \in \mathcal{V}(r)} \mu_z^*(r') \\ \mu_z^*(r') &= \delta(z(r') - z(r)) f(r') + (1 - \delta(z(r') - z(r))) m_z(r') \\ &= (1 - c(r')) f(r') + c(r') m_z(r')\end{aligned}$$

$$\begin{aligned}p(f|z) &\propto \prod_r \mathcal{N}(\mu_z(r), v_z(r)) \propto \prod_k \alpha_k \mathcal{N}(m_k 1, \Sigma_k) \\ p(z) &= \prod_r v_z(r) = \prod_k \alpha_k^{n_k}\end{aligned}$$

with  $1_k = 1, \forall r \in \mathcal{R}_k$  and  $\Sigma_k$  a covariance matrix  $(n_k \times n_k)$ .

## Case 3: $f|z$ Gauss iid, $z$ Potts

Gauss iid as in Case 1:

$$p(f|z) = \prod_{r \in \mathcal{R}} \mathcal{N}(m_z(r), v_z(r)) = \prod_k \prod_{r \in \mathcal{R}_k} \mathcal{N}(m_k, v_k)$$

Potts-Markov

$$p(z(r)|z(r'), r' \in \mathcal{V}(r)) \propto \exp \left\{ \gamma \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}$$

$$p(z) \propto \exp \left\{ \gamma \sum_{r \in \mathcal{R}} \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}$$

## Case 4: $f|z$ Gauss-Markov, $z$ Potts

Gauss-Markov as in Case 2:

$$p(f(\mathbf{r})|z(\mathbf{r}), z(\mathbf{r}'), f(\mathbf{r}'), \mathbf{r}' \in \mathcal{V}(\mathbf{r})) = \mathcal{N}(\mu_z(\mathbf{r}), v_z(\mathbf{r})), \forall \mathbf{r} \in \mathcal{R}$$

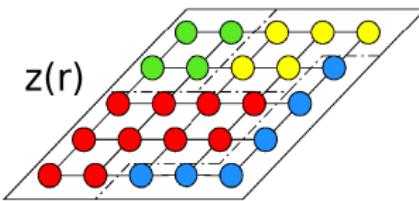
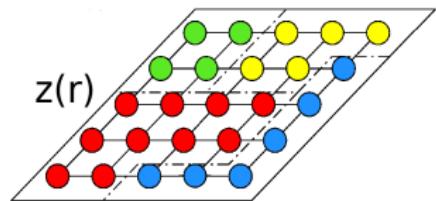
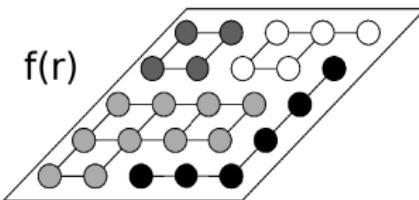
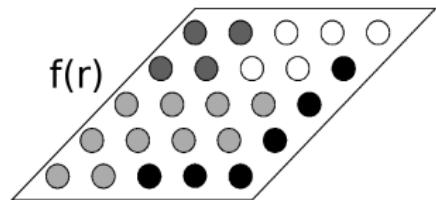
$$\begin{aligned}\mu_z(\mathbf{r}) &= \frac{1}{|\mathcal{V}(\mathbf{r})|} \sum_{\mathbf{r}' \in \mathcal{V}(\mathbf{r})} \mu_z^*(\mathbf{r}') \\ \mu_z^*(\mathbf{r}') &= \delta(z(\mathbf{r}') - z(\mathbf{r})) f(\mathbf{r}') + (1 - \delta(z(\mathbf{r}') - z(\mathbf{r}))) m_z(\mathbf{r}')\end{aligned}$$

$$p(f|z) \propto \prod_{\mathbf{r}} \mathcal{N}(\mu_z(\mathbf{r}), v_z(\mathbf{r})) \propto \prod_k \alpha_k \mathcal{N}(m_k \mathbf{1}, \boldsymbol{\Sigma}_k)$$

Potts-Markov as in Case 3:

$$p(z) \propto \exp \left\{ \gamma \sum_{\mathbf{r} \in \mathcal{R}} \sum_{\mathbf{r}' \in \mathcal{V}(\mathbf{r})} \delta(z(\mathbf{r}) - z(\mathbf{r}')) \right\}$$

## Summary of the two proposed models



$f|z$  Gaussian iid  
 $z$  Potts-Markov

$f|z$  Markov  
 $z$  Potts-Markov

(MIG with Hidden Potts)

(MGM with hidden Potts)

# Bayesian Computation

$$p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \mathbf{z}, v_\epsilon) p(\mathbf{f} | \mathbf{z}, \mathbf{m}, \mathbf{v}) p(\mathbf{z} | \gamma, \boldsymbol{\alpha}) p(\boldsymbol{\theta})$$

$$\boldsymbol{\theta} = \{v_\epsilon, (\alpha_k, m_k, v_k), k = 1, \dots, K\} \quad p(\boldsymbol{\theta}) \text{ Conjugate priors}$$

- ▶ Direct computation and use of  $p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g}; \mathcal{M})$  is too complex
- ▶ Possible approximations :
  - ▶ Gauss-Laplace (Gaussian approximation)
  - ▶ Exploration (Sampling) using MCMC methods
  - ▶ Separable approximation (Variational techniques)
- ▶ Main idea in Variational Bayesian methods:  
Approximate  
 $p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g}; \mathcal{M}) \quad \text{by} \quad q(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta}) = q_1(\mathbf{f}) q_2(\mathbf{z}) q_3(\boldsymbol{\theta})$ 
  - ▶ Choice of approximation criterion :  $KL(q : p)$
  - ▶ Choice of appropriate families of probability laws for  $q_1(\mathbf{f})$ ,  $q_2(\mathbf{z})$  and  $q_3(\boldsymbol{\theta})$

# MCMC based algorithm

$$p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \mathbf{z}, \boldsymbol{\theta}) p(\mathbf{f} | \mathbf{z}, \boldsymbol{\theta}) p(\mathbf{z}) p(\boldsymbol{\theta})$$

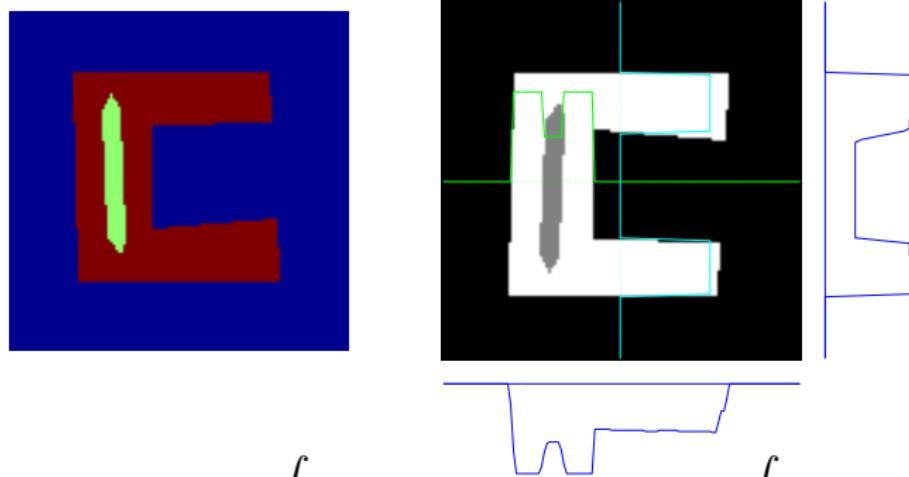
General scheme:

$$\hat{\mathbf{f}} \sim p(\mathbf{f} | \hat{\mathbf{z}}, \hat{\boldsymbol{\theta}}, \mathbf{g}) \longrightarrow \hat{\mathbf{z}} \sim p(\mathbf{z} | \hat{\mathbf{f}}, \hat{\boldsymbol{\theta}}, \mathbf{g}) \longrightarrow \hat{\boldsymbol{\theta}} \sim (\boldsymbol{\theta} | \hat{\mathbf{f}}, \hat{\mathbf{z}}, \mathbf{g})$$

- ▶ Estimate  $\mathbf{f}$  using  $p(\mathbf{f} | \hat{\mathbf{z}}, \hat{\boldsymbol{\theta}}, \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \boldsymbol{\theta}) p(\mathbf{f} | \hat{\mathbf{z}}, \hat{\boldsymbol{\theta}})$   
Needs optimisation of a quadratic criterion.
- ▶ Estimate  $\mathbf{z}$  using  $p(\mathbf{z} | \hat{\mathbf{f}}, \hat{\boldsymbol{\theta}}, \mathbf{g}) \propto p(\mathbf{g} | \hat{\mathbf{f}}, \hat{\mathbf{z}}, \hat{\boldsymbol{\theta}}) p(\mathbf{z})$   
Needs sampling of a Potts Markov field.
- ▶ Estimate  $\boldsymbol{\theta}$  using  
 $p(\boldsymbol{\theta} | \hat{\mathbf{f}}, \hat{\mathbf{z}}, \mathbf{g}) \propto p(\mathbf{g} | \hat{\mathbf{f}}, \sigma_\epsilon^2 \mathbf{I}) p(\hat{\mathbf{f}} | \hat{\mathbf{z}}, (m_k, v_k)) p(\boldsymbol{\theta})$   
Conjugate priors  $\longrightarrow$  analytical expressions.

# Application of CT in NDT

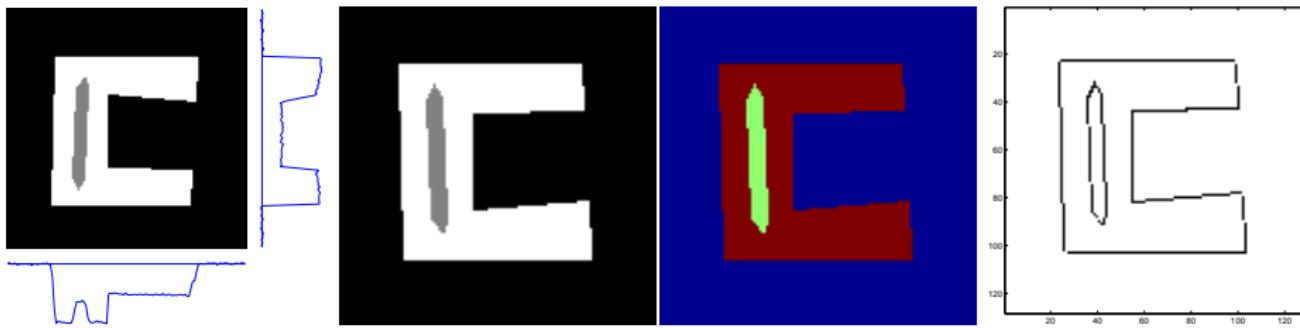
Reconstruction from only 2 projections



- Given the marginals  $g_1(x)$  and  $g_2(y)$  find the joint distribution  $f(x, y)$ .
- Infinite number of solutions :  $f(x, y) = g_1(x) g_2(y) \Omega(x, y)$   
 $\Omega(x, y)$  is a Copula:

$$\int \Omega(x, y) dx = 1 \quad \text{and} \quad \int \Omega(x, y) dy = 1$$

# Application in CT



$$\begin{aligned} \mathbf{g} | \mathbf{f} & \\ \mathbf{g} = \mathbf{H} \mathbf{f} + \epsilon & \\ \mathbf{g} | \mathbf{f} \sim \mathcal{N}(\mathbf{H} \mathbf{f}, \sigma_\epsilon^2 \mathbf{I}) & \\ \text{Gaussian} & \end{aligned}$$

$\mathbf{f} | \mathbf{z}$   
iid Gaussian  
or  
Gauss-Markov

$\mathbf{z}$   
iid  
or  
Potts

$\mathbf{c}$   
 $c(r) \in \{0, 1\}$   
 $1 - \delta(z(r) - z(r'))$   
binary

# Proposed algorithm

$$p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \mathbf{z}, \boldsymbol{\theta}) p(\mathbf{f} | \mathbf{z}, \boldsymbol{\theta}) p(\boldsymbol{\theta})$$

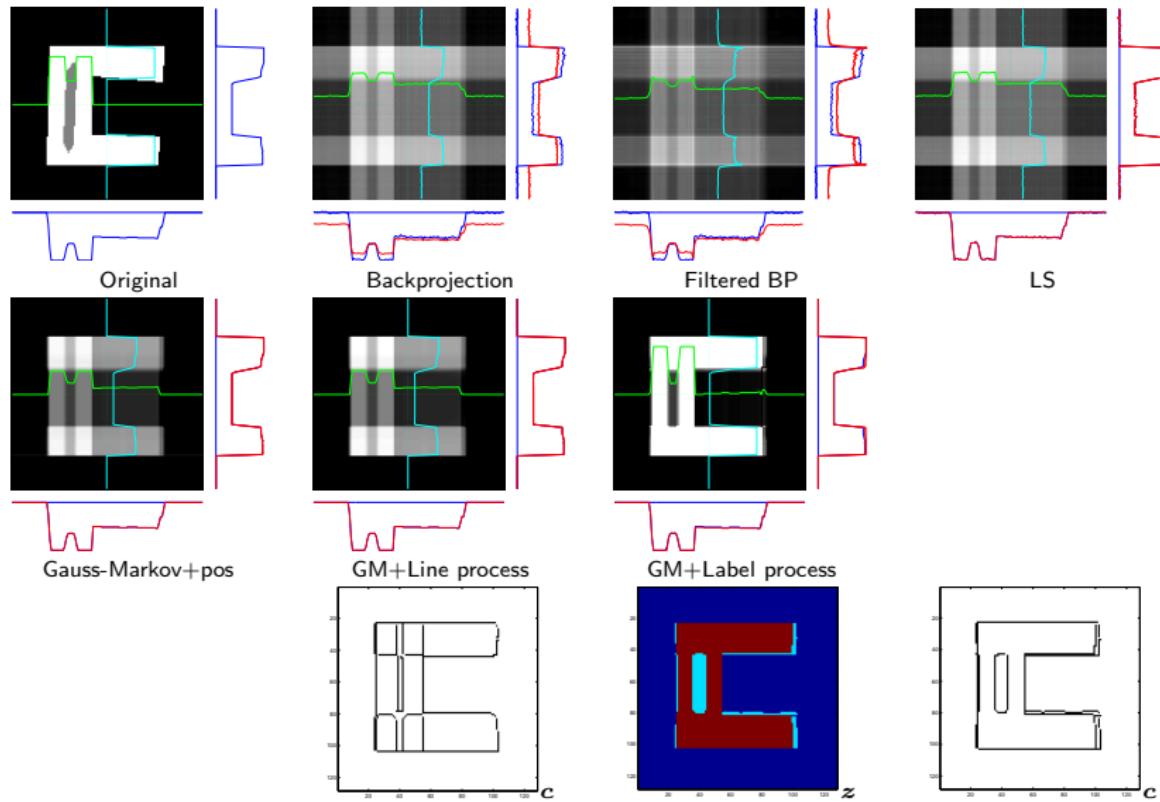
General scheme:

$$\hat{\mathbf{f}} \sim p(\mathbf{f} | \hat{\mathbf{z}}, \hat{\boldsymbol{\theta}}, \mathbf{g}) \longrightarrow \hat{\mathbf{z}} \sim p(\mathbf{z} | \hat{\mathbf{f}}, \hat{\boldsymbol{\theta}}, \mathbf{g}) \longrightarrow \hat{\boldsymbol{\theta}} \sim (\boldsymbol{\theta} | \hat{\mathbf{f}}, \hat{\mathbf{z}}, \mathbf{g})$$

Iterative algorithme:

- ▶ Estimate  $\mathbf{f}$  using  $p(\mathbf{f} | \hat{\mathbf{z}}, \hat{\boldsymbol{\theta}}, \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \boldsymbol{\theta}) p(\mathbf{f} | \hat{\mathbf{z}}, \hat{\boldsymbol{\theta}})$   
Needs **optimisation** of a quadratic criterion.
- ▶ Estimate  $\mathbf{z}$  using  $p(\mathbf{z} | \hat{\mathbf{f}}, \hat{\boldsymbol{\theta}}, \mathbf{g}) \propto p(\mathbf{g} | \hat{\mathbf{f}}, \hat{\mathbf{z}}, \hat{\boldsymbol{\theta}}) p(\mathbf{z})$   
Needs **sampling of a Potts Markov field**.
- ▶ Estimate  $\boldsymbol{\theta}$  using  
 $p(\boldsymbol{\theta} | \hat{\mathbf{f}}, \hat{\mathbf{z}}, \mathbf{g}) \propto p(\mathbf{g} | \hat{\mathbf{f}}, \sigma_\epsilon^2 \mathbf{I}) p(\hat{\mathbf{f}} | \hat{\mathbf{z}}, (m_k, v_k)) p(\boldsymbol{\theta})$   
Conjugate priors → **analytical expressions**.

# Results

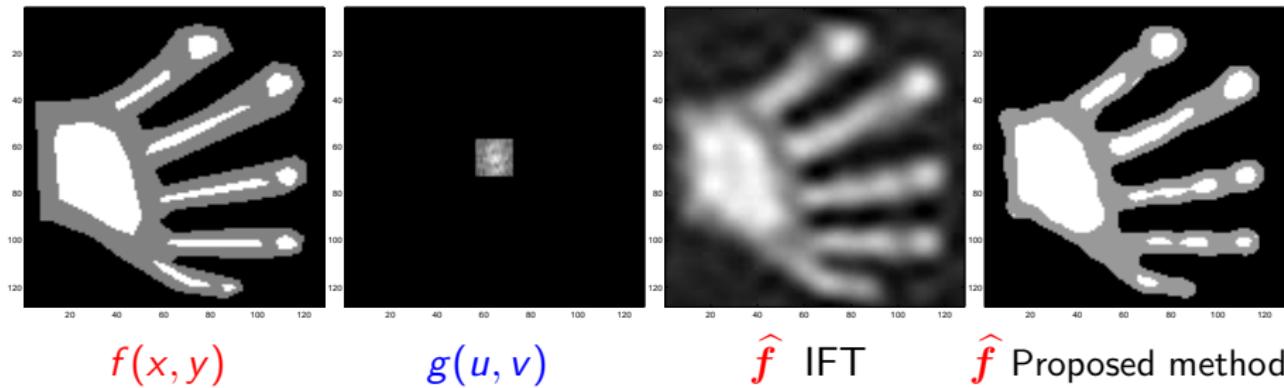


# Application in Microwave imaging

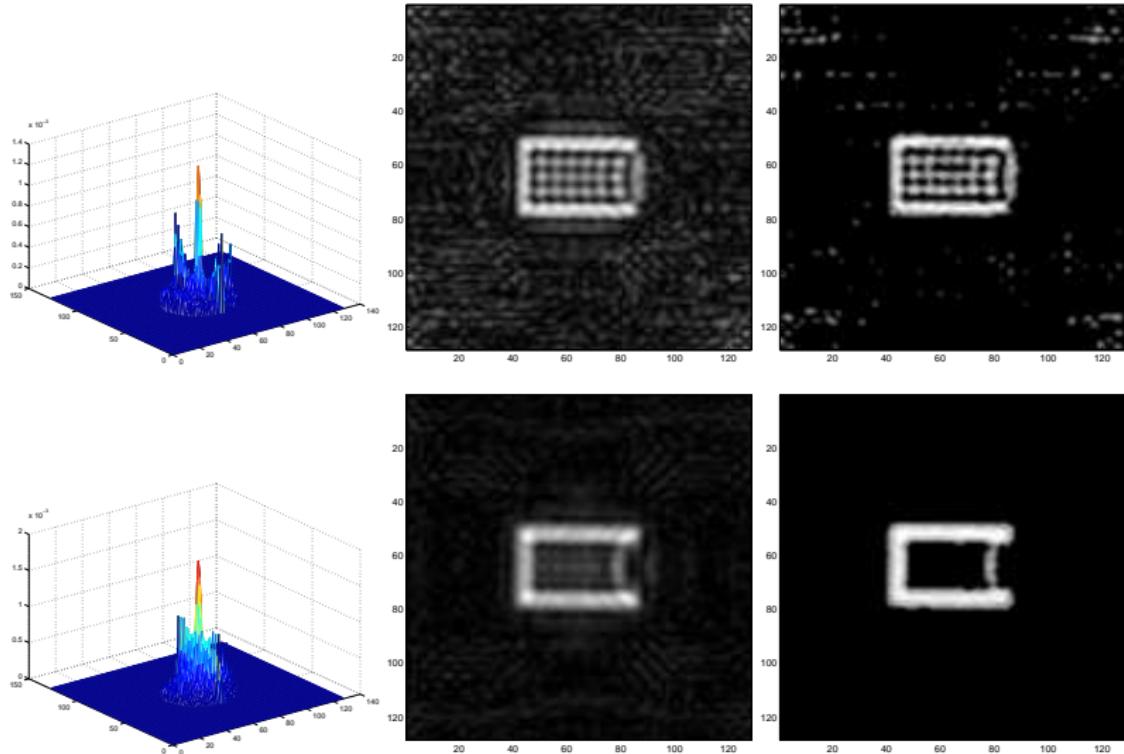
$$g(\omega) = \int f(r) \exp \{-j(\omega \cdot r)\} dr + \epsilon(\omega)$$

$$g(u, v) = \iint f(x, y) \exp \{-j(ux + vy)\} dx dy + \epsilon(u, v)$$

$$\mathbf{g} = \mathbf{H} \mathbf{f} + \boldsymbol{\epsilon}$$



# Application in Microwave imaging



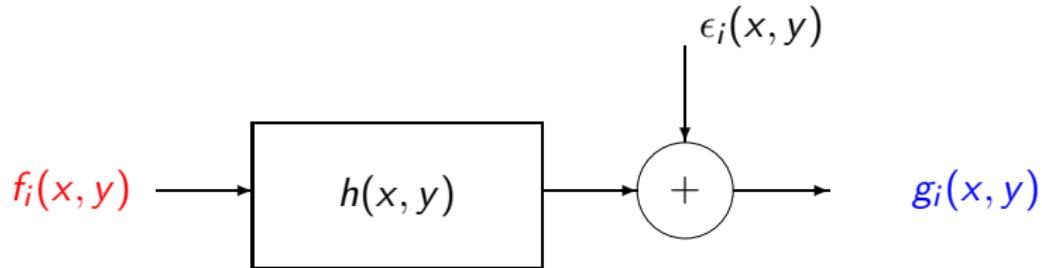
# Conclusions

- ▶ Bayesian Inference for inverse problems
- ▶ Approximations (Laplace, MCMC, Variational)
- ▶ Gauss-Markov-Potts are useful prior models for images incorporating regions and contours
- ▶ Separable approximations for Joint posterior with Gauss-Markov-Potts priors
- ▶ Application in different CT (X ray, US, Microwaves, PET, SPECT)

Perspectives :

- ▶ Efficient implementation in 2D and 3D cases
- ▶ Evaluation of performances and comparison with MCMC methods
- ▶ Application to other linear and non linear inverse problems: (PET, SPECT or ultrasound and microwave imaging)

# Color (Multi-spectral) image deconvolution



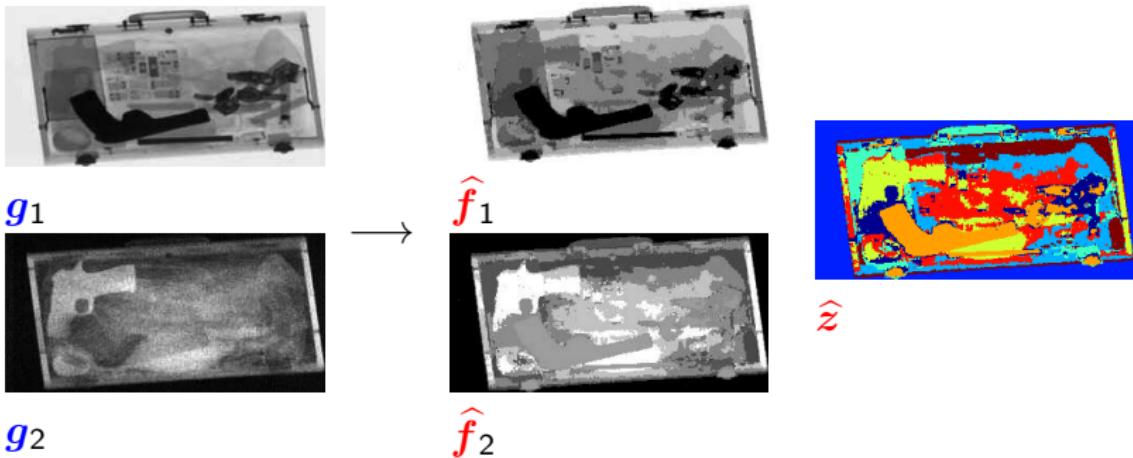
Observation model :  $\mathbf{g}_i = \mathbf{H} \mathbf{f}_i + \boldsymbol{\epsilon}_i, \quad i = 1, 2, 3$



# Images fusion and joint segmentation

(with O. Féron)

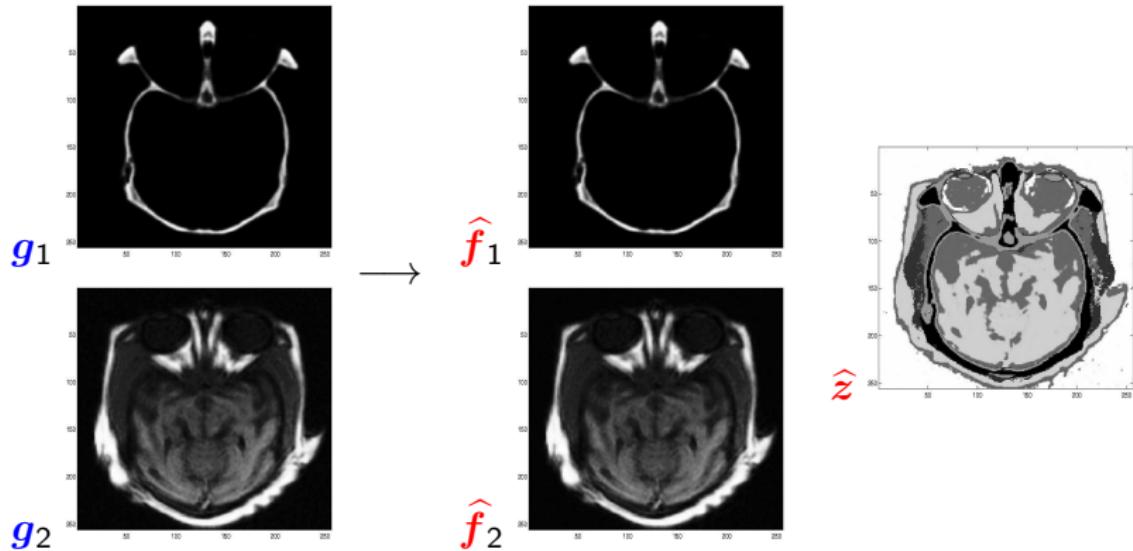
$$\begin{cases} \mathbf{g}_i(\mathbf{r}) = \mathbf{f}_i(\mathbf{r}) + \epsilon_i(\mathbf{r}) \\ p(f_i(\mathbf{r})|z(\mathbf{r}) = k) = \mathcal{N}(m_{ik}, \sigma_{ik}^2) \\ p(\underline{\mathbf{f}}|z) = \prod_i p(\mathbf{f}_i|z) \end{cases}$$



# Data fusion in medical imaging

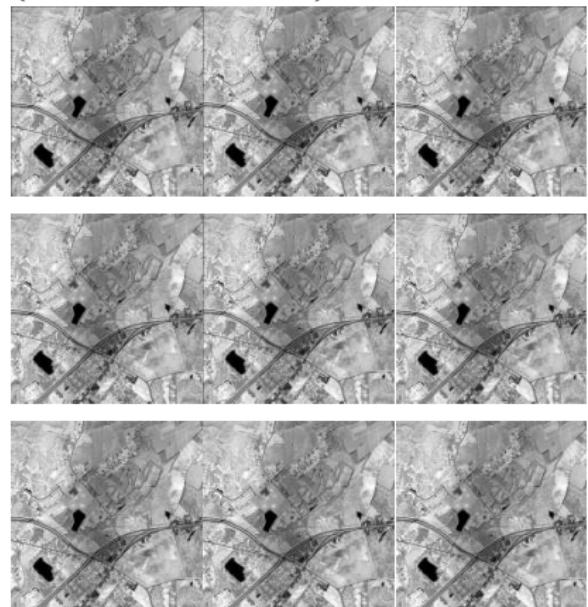
(with O. Féron)

$$\begin{cases} \underline{g_i(r)} = \underline{f_i(r)} + \epsilon_i(r) \\ p(f_i(r)|z(r) = k) = \mathcal{N}(m_{ik}, \sigma_{ik}^2) \\ p(\underline{f}|z) = \prod_i p(f_i|z) \end{cases}$$



# Super-Resolution

(with F. Humblot)



Low Resolution images

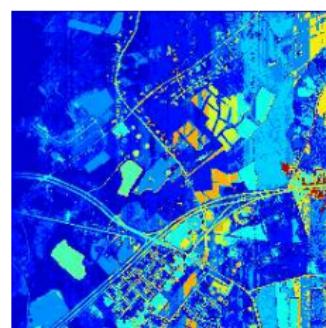
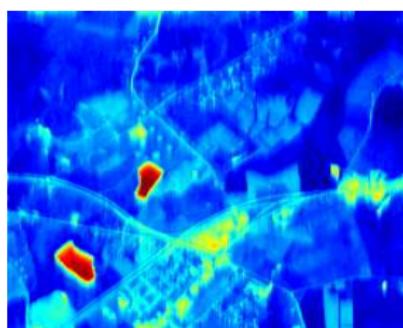
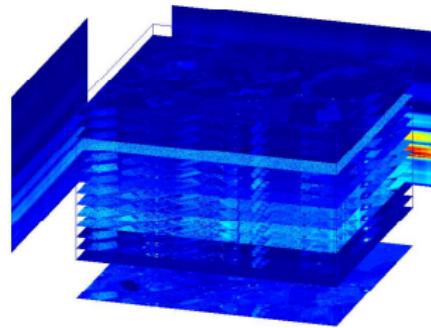


High Resolution image

# Joint segmentation of hyper-spectral images

(with N. Bali & A. Mohammadpour)

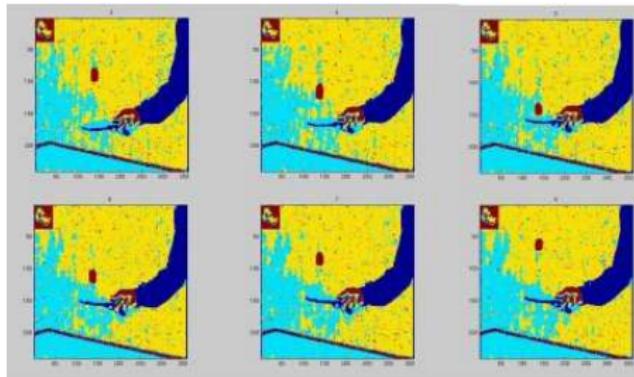
$$\begin{cases} g_i(\mathbf{r}) = f_i(\mathbf{r}) + \epsilon_i(\mathbf{r}) \\ p(f_i(\mathbf{r})|z(\mathbf{r}) = k) = \mathcal{N}(m_{ik}, \sigma_{ik}^2), \quad k = 1, \dots, K \\ p(\underline{\mathbf{f}}|\mathbf{z}) = \prod_i p(f_i|\mathbf{z}) \\ m_{ik} \text{ follow a Markovian model along the index } i \end{cases}$$



# Segmentation of a video sequence of images

(with P. Brault)

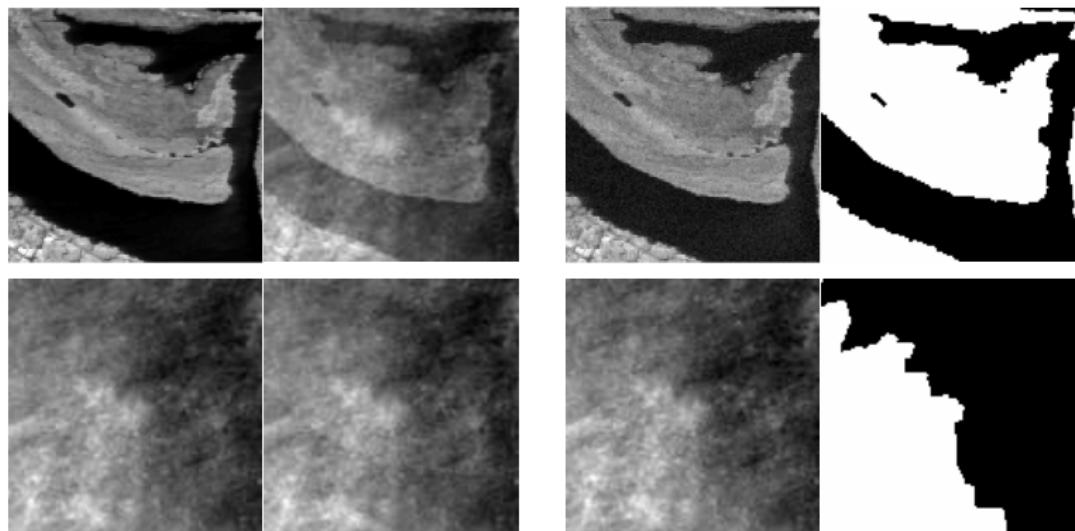
$$\begin{cases} g_i(\mathbf{r}) = f_i(\mathbf{r}) + \epsilon_i(\mathbf{r}) \\ p(f_i(\mathbf{r})|z_i(\mathbf{r}) = k) = \mathcal{N}(m_{ik}, \sigma_{ik}^2), \quad k = 1, \dots, K \\ p(\underline{\mathbf{f}}|\mathbf{z}) = \prod_i p(f_i|\mathbf{z}_i) \\ z_i(\mathbf{r}) \text{ follow a Markovian model along the index } i \end{cases}$$



# Source separation

(with H. Snoussi & M. Ichir)

$$\begin{cases} g_i(\mathbf{r}) = \sum_{j=1}^N A_{ij} f_j(\mathbf{r}) + \epsilon_i(\mathbf{r}) \\ p(f_j(\mathbf{r}) | z_j(\mathbf{r}) = k) = \mathcal{N}(m_{jk}, \sigma_{jk}^2) \\ p(A_{ij}) = \mathcal{N}(A_{0ij}, \sigma_{0ij}^2) \end{cases}$$



# Some references

- ▶ A. Mohammad-Djafari (Ed.) Problèmes inverses en imagerie et en vision (Vol. 1 et 2), [Hermes-Lavoisier, Traité Signal et Image, IC2](#), 2009,
- ▶ A. Mohammad-Djafari (Ed.) Inverse Problems in Vision and 3D Tomography, [ISTE, Wiley and sons](#), ISBN: [9781848211728](#), December 2009, Hardback, 480 pp.
- ▶ H. Ayasso and Ali Mohammad-Djafari Joint NDT Image Restoration and Segmentation using Gauss-Markov-Potts Prior Models and Variational Bayesian Computation, [To appear in IEEE Trans. on Image Processing, TIP-04815-2009.R2](#), 2010.
- ▶ H. Ayasso, B. Duchene and A. Mohammad-Djafari, Bayesian Inversion for Optical Diffraction Tomography [Journal of Modern Optics](#), 2008.
- ▶ A. Mohammad-Djafari, Gauss-Markov-Potts Priors for Images in Computer Tomography Resulting to Joint Optimal Reconstruction and segmentation, [International Journal of Tomography & Statistics 11: W09. 76-92](#), 2008.
- ▶ A Mohammad-Djafari, Super-Resolution : A short review, a new method based on hidden Markov modeling of HR image and future challenges, [The Computer Journal doi:10.1093/comjnl/bxn005](#), 2008.
- ▶ O. Féron, B. Duchène and A. Mohammad-Djafari, Microwave imaging of inhomogeneous objects made of a finite number of dielectric and conductive materials from experimental data, [Inverse Problems](#), 21(6):95-115, Dec 2005.
- ▶ M. Ichir and A. Mohammad-Djafari, Hidden markov models for blind source separation, [IEEE Trans. on Signal Processing](#), 15(7):1887-1899, Jul 2006.
- ▶ F. Humblot and A. Mohammad-Djafari, Super-Resolution using Hidden Markov Model and Bayesian Detection Estimation Framework, [EURASIP Journal on Applied Signal Processing, Special number on Super-Resolution Imaging: Analysis, Algorithms, and Applications:ID 36971, 16 pages](#), 2006.
- ▶ O. Féron and A. Mohammad-Djafari, Image fusion and joint segmentation using an MCMC algorithm, [Journal of Electronic Imaging](#), 14(2):paper no. 023014, Apr 2005.
- ▶ H. Snoussi and A. Mohammad-Djafari, Fast joint separation and segmentation of mixed images, [Journal of Electronic Imaging](#), 13(2):349-361, April 2004.
- ▶ A. Mohammad-Djafari, J.F. Giovannelli, G. Demoment and J. Idier, Regularization, maximum entropy and probabilistic methods in mass spectrometry data processing problems, [Int. Journal of Mass Spectrometry](#), 215(1-3):175-193, April 2002.

# Thanks, Questions and Discussions

## Thanks to:

### My graduated PhD students:

- ▶ H. Snoussi, M. Ichir, (Sources separation)
- ▶ F. Humblot (Super-resolution)
- ▶ H. Carfantan, O. Féron (Microwave Tomography)
- ▶ S. Fékih-Salem (3D X ray Tomography)

### My present PhD students:

- ▶ H. Ayasso (Optical Tomography, Variational Bayes)
- ▶ D. Pougaza (Tomography and Copula)
- ▶ \_\_\_\_\_
- ▶ Sh. Zhu (SAR Imaging)
- ▶ D. Fall (Emission Positon Tomography, Non Parametric Bayesian)

### My colleagues in GPI (L2S) & collaborators in other instituts:

- ▶ B. Duchêne & A. Joisel (Inverse scattering and Microwave Imaging)
- ▶ N. Gac & A. Rabanal (GPU Implementation)
- ▶ Th. Rodet (Tomography)
- ▶ \_\_\_\_\_
- ▶ A. Vabre & S. Legoupil (CEA-LIST), (3D X ray Tomography)
- ▶ E. Barat (CEA-LIST) (Positon Emission Tomography, Non Parametric Bayesian)
- ▶ C. Comtat (SHFJ, CEA)(PET, Spatio-Temporal Brain activity)

## Questions and Discussions