

# Sparsity in signal and image processing: from modeling and representation to reconstruction and processing

Ali Mohammad-Djafari

Laboratoire des Signaux et Systèmes,  
UMR8506 CNRS-SUPELEC-UNIV PARIS SUD  
SUPELEC, 91192 Gif-sur-Yvette, France

<http://lss.supelec.free.fr>

Email: [djafari@lss.supelec.fr](mailto:djafari@lss.supelec.fr)

<http://djafari.free.fr>

<http://publicationslist.org/djafari>

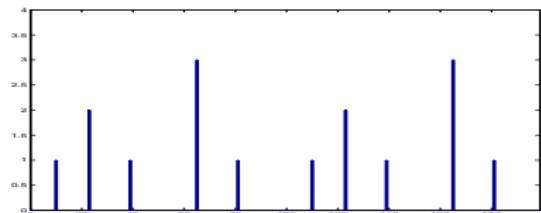
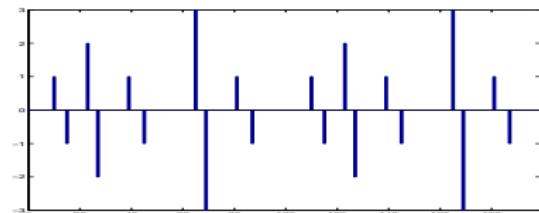
Seminar 2 given at  
Tsinghua University, Beijing, China, December 10, 2013

# Contents

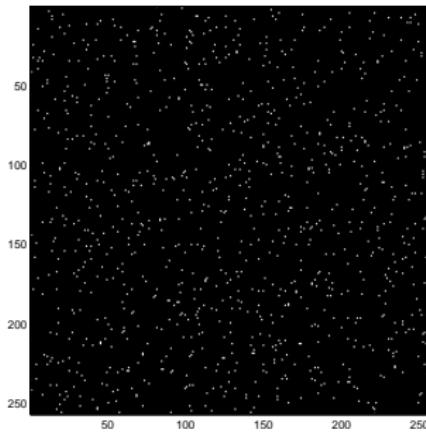
1. Sparse signals and images
2. First ideas for using sparsity in signal processing
3. Modeling for sparse representation
4. Bayesian Maximum A Posteriori (MAP) approach and link with Deterministic Regularization
5. Priors which enforce sparsity
  - ▶ Heavy tailed: Double Exponential, Generalized Gaussian, ...
  - ▶ Mixture models: Mixture of Gaussians, Student-t, ...
  - ▶ Hierarchical models with hidden variables
  - ▶ General Gauss-Markov-Potts models
6. Computational tools:  
Joint Maximum A Posteriori (JMAP), MCMC and Variational Bayesian Approximation (VBA)
7. Applications:  
X ray Computed Tomography, Microwave and Ultrasound imaging, Sattelite and Hyperspectral image processing, Spectrometry, CMB, ...

# 1. Sparse signals and images

- Sparse signals: Direct sparsity

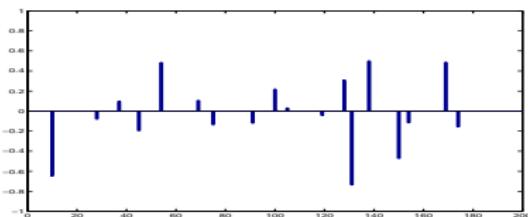
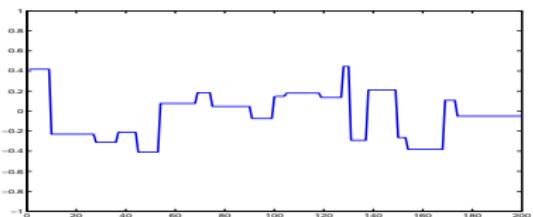


- Sparse images: Direct sparsity

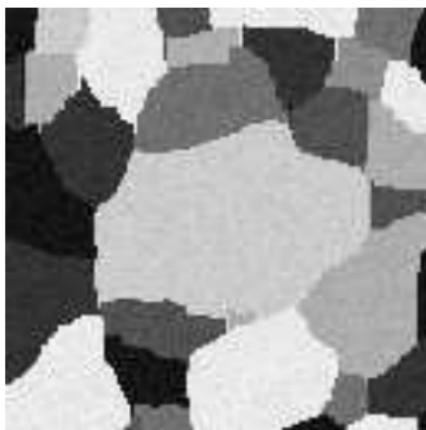


# Sparse signals and images

- ▶ Sparse signals in a Transform domain



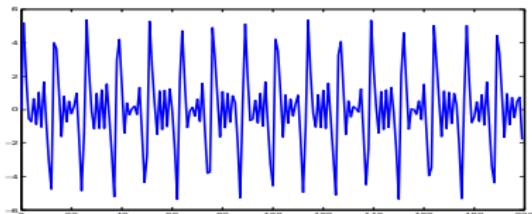
- ▶ Sparse images in a Transform domain



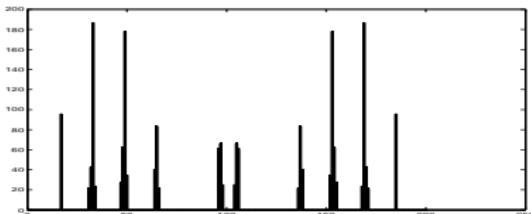
# Sparse signals and images

- ▶ Sparse signals in Fourier domain

Time domain



Fourier domain

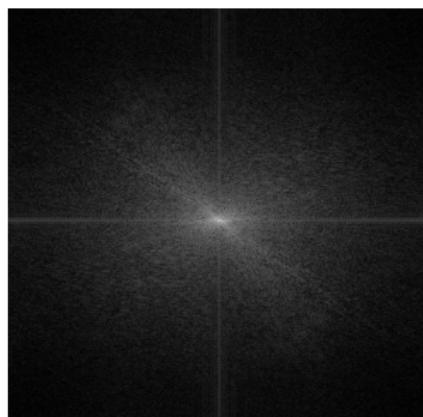


- ▶ Sparse images in wavelet domain

Space domain

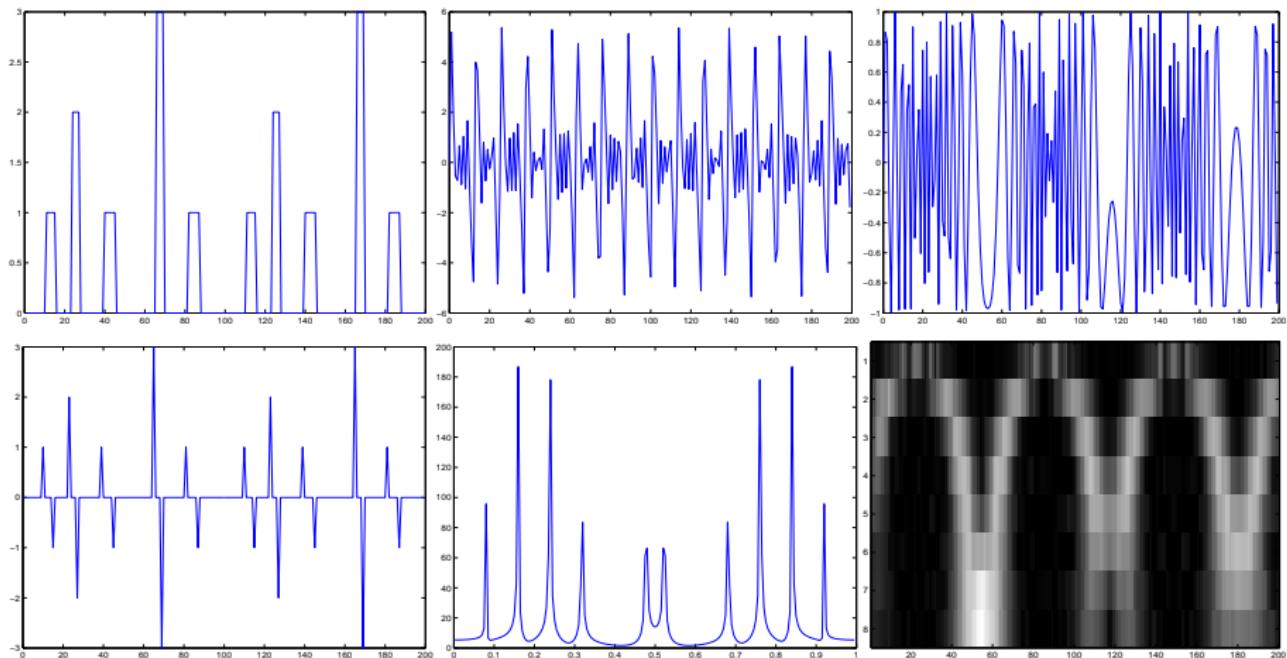


Fourier domain

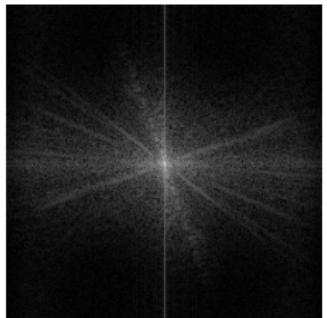
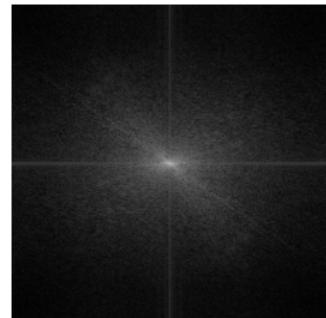
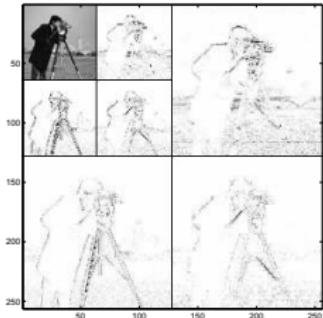
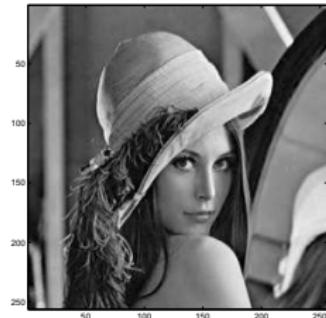


# Sparse signals and images

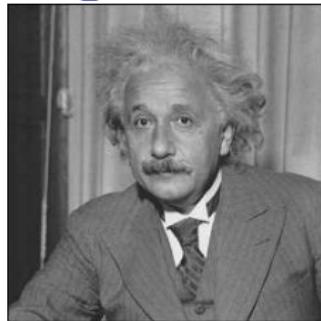
- Sparse signals: Sparsity in a Transform domain



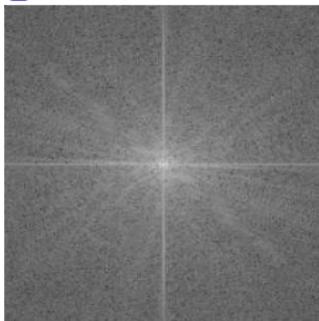
# Sparse signals and images



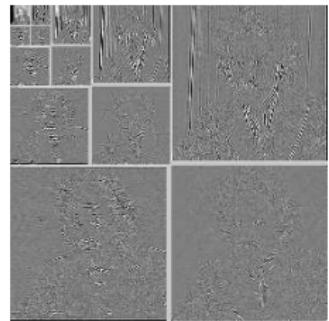
# Sparse signals and images



Image



Fourier



Wavelets

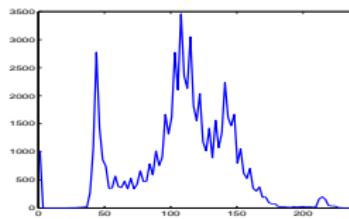
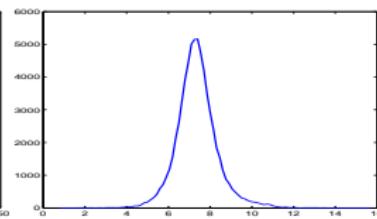
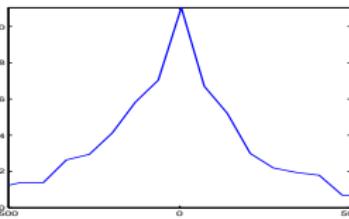


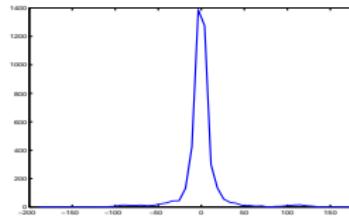
Image hist.



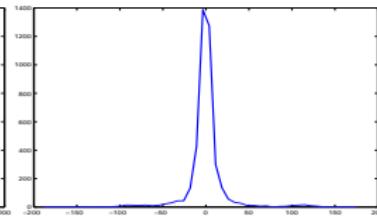
Fourier coeff. hist.



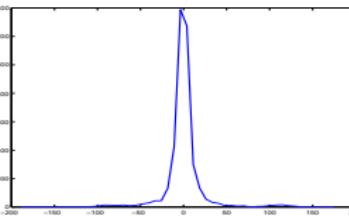
Wavelet coeff. hist.



bands 1-3



bands 4-6



bands 7-9

## 2. First ideas: some history

- ▶ 1948: Shannon:  
Sampling theorem and reconstruction of a band limited signal
- ▶ 1993-2007:
  - ▶ Mallat, Zhang, Candès, Romberg, Tao and Baraniuk:  
Non linear sampling, Compression and reconstruction,
  - ▶ Fuchs: Sparse representation
  - ▶ Donoho, Elad, Tibshirani, Tropp, Duarte, Laska:  
Compressive Sampling, Compressive Sensing
- ▶ 2007-2012:  
Algorithms for sparse representation and compressive Sampling: Matching Pursuit (MP), Projection Pursuit Regression, Pure Greedy Algorithm, OMP, Basis Pursuit (BP), Dantzig Selector (DS), Least Absolute Shrinkage and Selection Operator (LASSO),...
- ▶ 2003-2012:  
Bayesian approach to sparse modeling  
Tipping, Bishop: Sparse Bayesian Learning,  
Relevance Vector Machine (RVM), ...

### 3. Modeling and representation

- Modeling via a basis (codebook, overcomplete dictionary, Design Matrix)

$$\mathbf{g}(t) = \sum_{j=1}^N \mathbf{f}_j \phi_j(t), \quad t = 1, \dots, T \longrightarrow \mathbf{g} = \Phi' \mathbf{f}$$

- When  $T \geq N$

$$\widehat{\mathbf{f}}_j = \arg \min_{f_j} \left\{ \sum_{t=1}^T \left| \mathbf{g}(t) - \sum_{j=1}^N \mathbf{f}_j \phi_j(t) \right|^2 \right\} \rightarrow$$
$$\widehat{\mathbf{f}} = \arg \min_{\mathbf{f}} \left\{ \|\mathbf{g} - \Phi' \mathbf{f}\|^2 \right\} = [\Phi \Phi']^{-1} \Phi \mathbf{g}$$

- When orthogonal basis:  $\Phi \Phi' = I \longrightarrow \widehat{\mathbf{f}} = \Phi \mathbf{g}$

$$\widehat{\mathbf{f}}_j = \sum_{t=1}^N \mathbf{g}(t) \phi_j(t) = \langle \mathbf{g}(t), \phi_j(t) \rangle$$

- Application in **Compression**, **Transmission** and **Decompression**

## Modeling and representation

- When overcomplete basis  $N > T$ : Infinite number of solutions for  $\Phi' \mathbf{f} = \mathbf{g}$ . We have to select one:

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}: \Phi' \mathbf{f} = \mathbf{g}} \{ \|\mathbf{f}\|_2^2 \}$$

or writing differently:

$$\text{minimize } \|\mathbf{f}\|_2^2 \text{ subject to } \Phi' \mathbf{f} = \mathbf{g}$$

resulting to:

$$\hat{\mathbf{f}} = \Phi [\Phi' \Phi]^{-1} \mathbf{g}$$

- Again if  $\Phi' \Phi = I \rightarrow \hat{\mathbf{f}} = \Phi \mathbf{g}$ .
- No real interest if we have to keep all the  $N > T$  coefficients.
- Sparsity:

$$\text{minimize } \|\mathbf{f}\|_0 \text{ subject to } \Phi' \mathbf{f} = \mathbf{g}$$

or

$$\text{minimize } \|\mathbf{f}\|_1 \text{ subject to } \Phi' \mathbf{f} = \mathbf{g}$$

# Sparse decomposition

- ▶ Strict sparsity and exact reconstruction

$$\text{minimize } \|\mathbf{f}\|_0 \text{ subject to } \Phi' \mathbf{f} = \mathbf{g}$$

$\|\mathbf{f}\|_0$  is the number of non-zero elements of  $\mathbf{f}$

- ▶ Matching Pursuit (MP) [Mallat & Zhang, 1993]
- ▶ Orthogonal Matching Pursuit (OMP) [Lin, Huang et al., 1993]
- ▶ Projection Pursuit Regression
- ▶ Greedy Algorithms

- ▶ Sparsity enforcing and exact reconstruction

$$\text{minimize } \|\mathbf{f}\|_1 \text{ subject to } \Phi' \mathbf{f} = \mathbf{g}$$

- ▶ Basis Pursuit (BP)
- ▶ Block Coordinate Relaxation

# Sparse decomposition

- ▶ Strict sparsity and approximate reconstruction

$$\text{minimize } \|\mathbf{f}\|_0 \text{ subject to } \|\mathbf{g} - \Phi' \mathbf{f}\|^2 < c$$

$$\widehat{\mathbf{f}} = \arg \min_{\mathbf{f}} \left\{ \|\mathbf{f}\|_0 + \mu \|\mathbf{g} - \Phi' \mathbf{f}\|^2 \right\} = \arg \min_{\mathbf{f}} \left\{ \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda \|\mathbf{f}\|_0 \right\}$$

- ▶ Sparsity enforcing and approximate reconstruction

$$\widehat{\mathbf{f}} = \arg \min_{\mathbf{f}} \left\{ \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda \|\mathbf{f}\|_1 \right\}$$

$$J(\mathbf{f}) = \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda \|\mathbf{f}\|_1 = \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda \sum_j |\mathbf{f}_j|$$

- ▶ Main Algorithm: LASSO [Tibshirani 2003]

$$\text{minimize } \|\mathbf{g} - \Phi' \mathbf{f}\|^2 \text{ subject to } \|\mathbf{f}\|_1 < \tau$$

# Sparse Decomposition Algorithms

- ▶ LASSO:

$$J(\mathbf{f}) = \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda \sum_j |\mathbf{f}_j|$$

- ▶ Other Criteria

- ▶  $L_p$

$$J(\mathbf{f}) = \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda_1 \sum_j |\mathbf{f}_j|^p, \quad 1 < p \leq 2$$

- ▶ Elastic net

$$J(\mathbf{f}) = \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \sum_j (\lambda_1 |\mathbf{f}_j| + \lambda_2 |\mathbf{f}_j|^2)$$

- ▶ Group LASSO

$$J(\mathbf{f}) = \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda_1 \sum_j |\mathbf{f}_j| + \lambda_2 \sum_j |\mathbf{f}_j - \mathbf{f}_{j-1}|^2$$

- ▶ Weighted L1:

$$J(\mathbf{f}) = \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda \sum_j |w_j \mathbf{f}_j|$$

# Multi Dimensional signals: PCA, SPCA, BSS, ...

$$\mathbf{g}_i(t) = \sum_{j=1}^N \Phi_{ij} \mathbf{f}_j(t), \quad i = 1, \dots, M, \quad t = 1, \dots, T$$

$$\mathbf{g}(t) = \Phi \mathbf{f}(t), \quad t = 1, \dots, T$$

$$\mathbf{G} = \Phi \mathbf{F}, \quad \text{with } \mathbf{G} [M \times T], \quad \Phi [M \times N], \quad \mathbf{F} [N \times T]$$

- ▶  $\mathbf{f}_j(t)$  factors, sources, codes
- ▶  $\Phi$  Loading matrix (Factor Analysis),  
Mixing matrix (Blind Sources Separation),  
Design matrix (Sparse coding, Compressed Sensing)
- ▶ Objective: Find  $\Phi$  and  $\mathbf{f}_j(t)$

$$J(\mathbf{f}(t), \Phi) = \sum_t \|\mathbf{g}(t) - \Phi \mathbf{f}(t)\|^2 + \lambda_1 \sum_i \sum_j |\Phi_{ij}| + \lambda_2 \sum_t \sum_j |\mathbf{f}_j(t)|$$

$$J(\mathbf{F}, \Phi) = \|\mathbf{G} - \Phi \mathbf{F}\|^2 + \lambda_1 \|\Phi\|_1 + \lambda_2 \|\mathbf{F}\|_1$$

# Matrix Decomposition or Approximation

- Matrix approximation:

Find an approximate matrix  $\hat{\mathbf{G}} = \Phi \mathbf{F}$  for  $\mathbf{G}$  with some degrees of sparsity in the elements of  $\Phi$  and  $\mathbf{F}$ .

$$J(\mathbf{F}, \Phi) = \|\mathbf{G} - \Phi \mathbf{F}\|^2 + \lambda_1 \|\Phi\|_1 + \lambda_2 \|\mathbf{F}\|_1$$

- Low rank Matrice decomposition:

$$\hat{\mathbf{G}} = \sum_{k=1}^K d_k \mathbf{u}_k \mathbf{v}'_k = \mathbf{U} \mathbf{D} \mathbf{V}$$

with some degrees of sparsity in the elements of  $\mathbf{u}$  and  $\mathbf{v}$ .

$$J(\mathbf{U}, \mathbf{V}) = \|\mathbf{G} - \mathbf{U} \mathbf{D} \mathbf{V}\|^2 + \lambda_1 \|\mathbf{U}\|_1 + \lambda_2 \|\mathbf{V}\|_1$$

## 4. Sparse decomposition: Bayesian MAP interpretation

- ▶ Sparsity enforcing and approximate reconstruction

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \left\{ \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda \|\mathbf{f}\|_1 \right\}$$

- ▶ Equivalent to MAP estimate in a Bayesian approach
  - ▶ Bayesian approach:

$$p(\mathbf{f}|\mathbf{g}) = \frac{p(\mathbf{g}|\mathbf{f}) p(\mathbf{f})}{p(\mathbf{g})} \propto p(\mathbf{g}|\mathbf{f}) p(\mathbf{f})$$

$$\begin{cases} p(\mathbf{g}|\mathbf{f}) = \mathcal{N}(\Phi' \mathbf{f}, \sigma_\epsilon^2) \propto \exp \left[ \frac{-1}{2\sigma_\epsilon^2} \|\mathbf{g} - \Phi' \mathbf{f}\|^2 \right] \\ p(\mathbf{f}) = \mathcal{D}\mathcal{E}(-\gamma) \propto \exp [\gamma \|\mathbf{f}\|_1] \end{cases}$$

- ▶ Maximum A Posteriori (MAP):

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} \{p(\mathbf{f}|\mathbf{g})\} = \arg \min_{\mathbf{f}} \{J(\mathbf{f})\}$$

$$J(\mathbf{f}) = \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda \|\mathbf{f}\|_1 \quad \text{with} \quad \lambda = 2\gamma\sigma_\epsilon^2$$

# Sparse decomposition: Regularization or MAP

- ▶ Regularization:  $J(\mathbf{f}) = \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda \|\mathbf{f}\|_1$
- ▶ With fixed  $\lambda$ : Find a good optimization algorithm
- ▶ How to choose  $\lambda$  ?
- ▶ L-Curve, Cross Validation, adhoc  $\lambda = 1, \dots$
- ▶ \_\_\_\_\_
- ▶ MAP:  $J(\mathbf{f}) = \|\mathbf{g} - \Phi' \mathbf{f}\|^2 + \lambda \|\mathbf{f}\|_1$  with  $\lambda = 2\gamma\sigma_\epsilon^2$
- ▶ How to estimate  $\gamma$  and  $\sigma_\epsilon^2$ ?
- ▶ Bayesian: Joint MAP, Expectation-Maximization, MCMC, Variational Bayesian Approximation,...
- ▶ Advantages of the Bayesian approach:
  - ▶ More probabilistic modeling for sparsity enforcing
  - ▶ Hyperparameter estimation
  - ▶ Uncertainty handling

## 5. Sparsity enforcing prior models

- ▶ Simple heavy tailed models:
  - ▶ Generalized Gaussian, Double Exponential
  - ▶ Student-t, Cauchy
  - ▶ Elastic net
  - ▶ Generalized hyperbolic
  - ▶ Symmetric Weibull, Symmetric Rayleigh
- ▶ Hierarchical mixture models:
  - ▶ Mixture of Gaussians
  - ▶ Bernoulli-Gaussian
  - ▶ Mixture of Gammas
  - ▶ Bernoulli-Gamma
  - ▶ Mixture of Dirichlet
  - ▶ Bernoulli-Multinomial

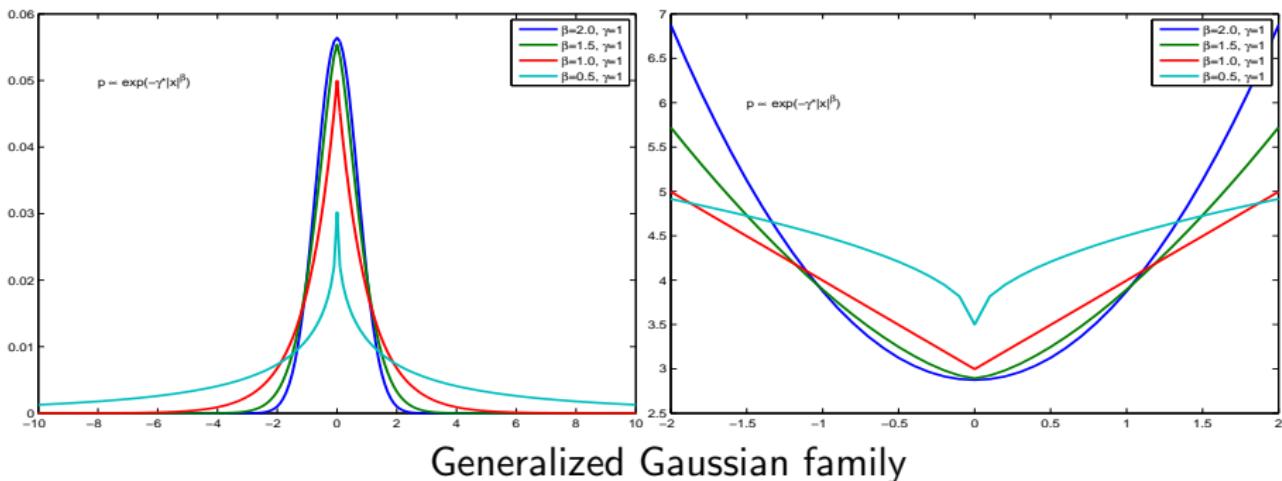
# Simple heavy tailed models

- Generalized Gaussian, Double Exponential

$$p(\mathbf{f}|\gamma, \beta) = \prod_j \mathcal{GG}(f_j|\gamma, \beta) \propto \exp \left[ -\gamma \sum_j |f_j|^\beta \right]$$

$\beta = 1$  Double exponential or Laplace.

$0 < \beta < 2$  are of great interest for sparsity enforcing.

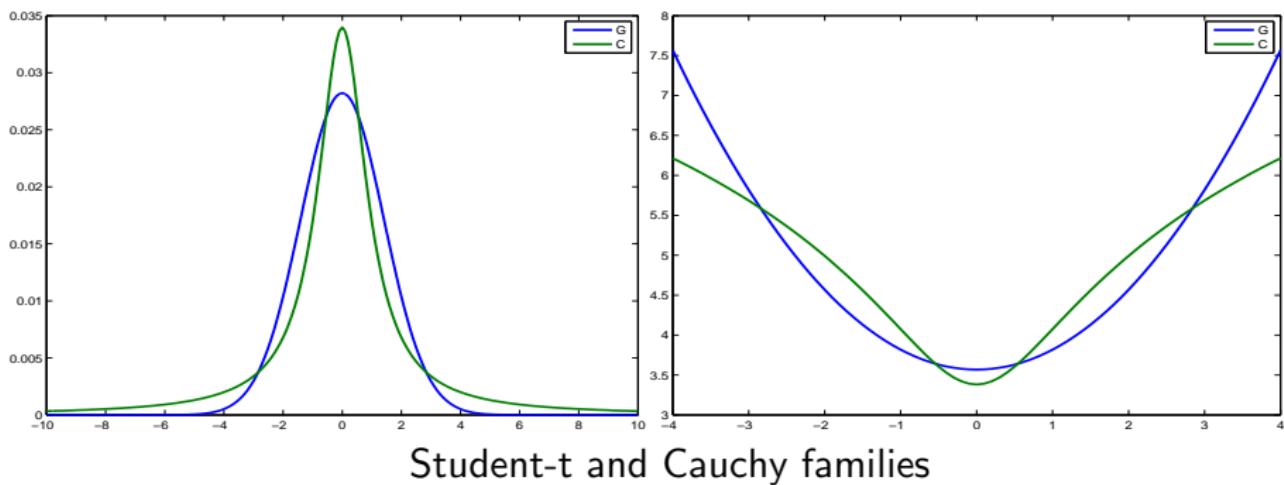


# Simple heavy tailed models

- Student-t and Cauchy models

$$p(\mathbf{f}|\nu) = \prod_j \mathcal{St}(f_j|\nu) \propto \exp \left[ -\frac{\nu+1}{2} \sum_j \log (1 + f_j^2/\nu) \right]$$

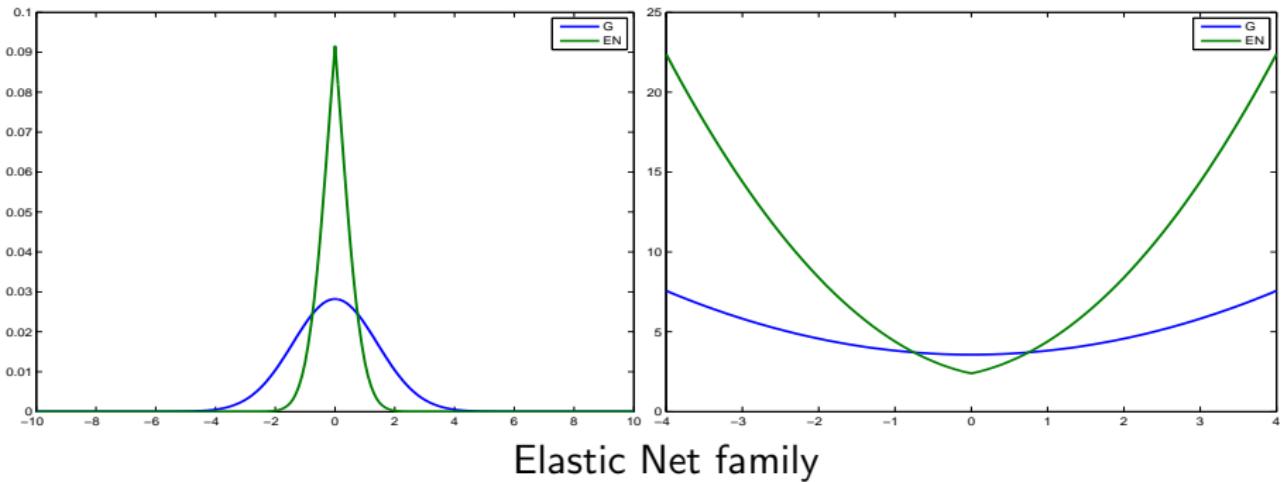
Cauchy model is obtained when  $\nu = 1$ .



# Simple heavy tailed models

- Elastic net prior model

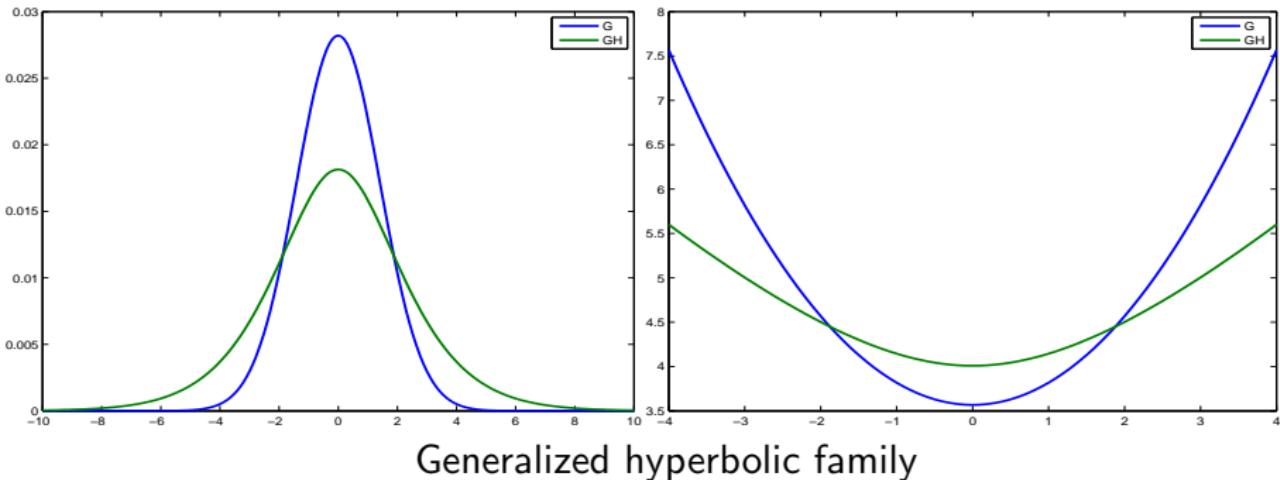
$$p(\mathbf{f}|\nu) = \prod_j \mathcal{EN}(f_j|\nu) \propto \exp \left[ - \sum_j (\gamma_1 |f_j| + \gamma_2 f_j^2) \right]$$



# Simple heavy tailed models

- Generalized hyperbolic (GH) models

$$p(\mathbf{f}|\delta, \nu, \beta) = \prod_j (\delta^2 + f_j^2)^{(\nu-1/2)/2} \exp [\beta x] K_{\nu-1/2}(\alpha \sqrt{\delta^2 + f_j^2})$$



# Simple heavy tailed models

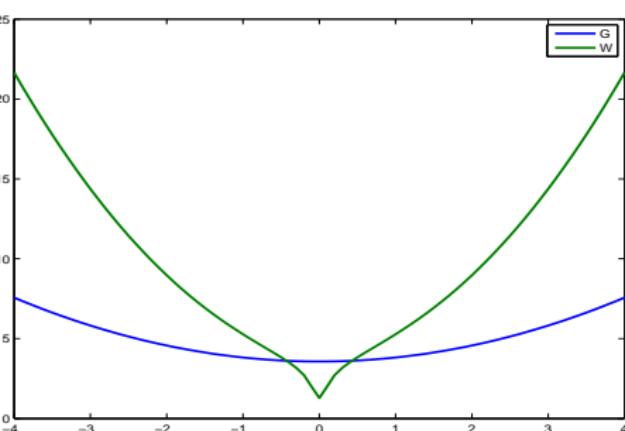
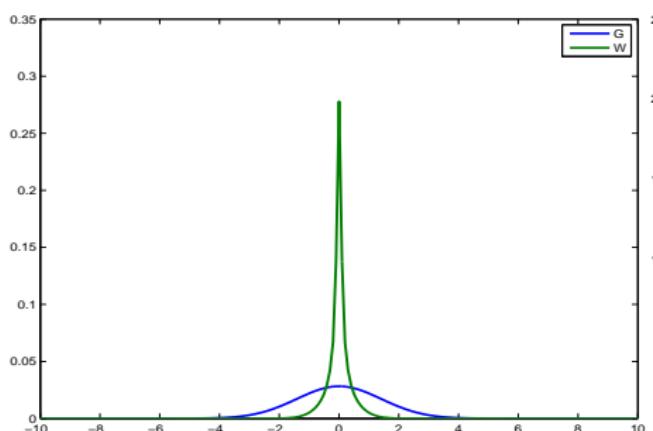
- Symmetric Weibull

$$p(\mathbf{f}|\gamma, \beta) = \prod_j \mathcal{W}(f_j|\gamma, \beta) \propto \exp \left[ -\gamma \sum_j |f_j|^\beta + (\beta - 1) \log |f_j| \right]$$

$\beta = 2$  is the Symmetric Rayleigh distribution.

$\beta = 1$  is the Double exponential and

$0 < \beta < 2$  are of great interest for sparsity enforcing.



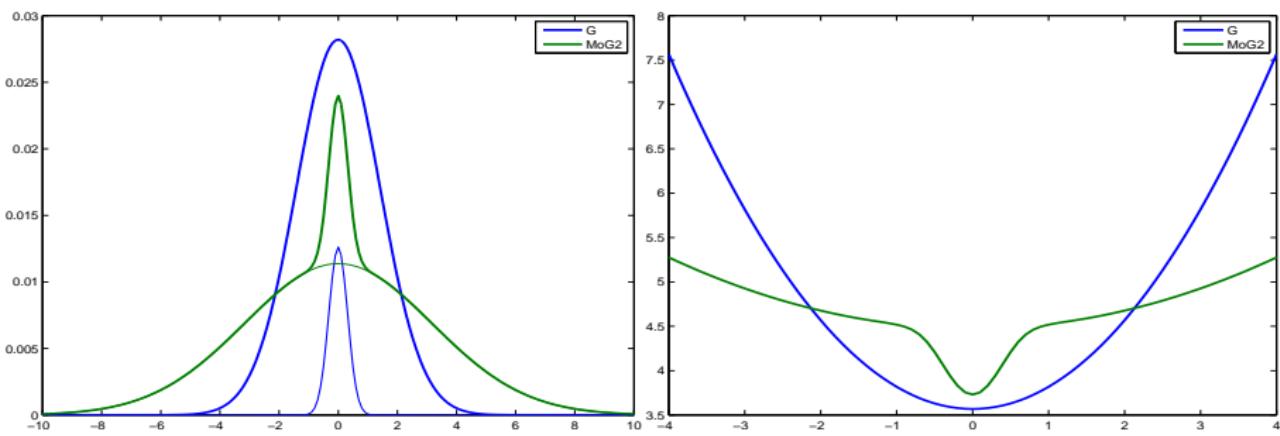
# Mixture models

- Mixture of two Gaussians (MoG2) model

$$p(\mathbf{f}|\alpha, v_1, v_0) = \prod_j [\alpha \mathcal{N}(f_j|0, v_1) + (1 - \alpha) \mathcal{N}(f_j|0, v_0)]$$

- Bernoulli-Gaussian (BG) model

$$p(\mathbf{f}|\alpha, v) = \prod_j p(f_j) = \prod_j [\alpha \mathcal{N}(f_j|0, v) + (1 - \alpha) \delta(f_j)]$$



Mixture of 2 Gaussians families

- Mixture of Gammas

$$p(\mathbf{f}|\lambda, v_1, v_0) = \prod_j [\lambda \mathcal{G}(f_j|\alpha_1, \beta_1) + (1 - \lambda) \mathcal{G}(f_j|\alpha_2, \beta_2)]$$

- Bernoulli-Gamma model

$$p(\mathbf{f}|\lambda, \alpha, \beta) = \prod_j [\lambda \mathcal{G}(f_j|\alpha, \beta) + (1 - \lambda) \delta(f_j)]$$

- Mixture of Dirichlets model

$$p(\mathbf{f}|\lambda, \mathbf{H}_1, \boldsymbol{\alpha}_1, \mathbf{H}_2, \boldsymbol{\alpha}_2) = \prod_j [\lambda \mathcal{D}(f_j|\mathbf{H}_1, \boldsymbol{\alpha}_1) + (1 - \lambda) \mathcal{D}(f_j|\mathbf{H}_2, \boldsymbol{\alpha}_2)]$$

$$\mathcal{D}(f_j|\mathbf{H}, \boldsymbol{\alpha}) = \prod_{k=1}^K \frac{\Gamma(\alpha)}{\Gamma(\alpha_0)\Gamma(\alpha_K)} a_k^{\alpha_k-1}, \quad \alpha_k \geq 0, \quad a_k \geq 0$$

where  $\mathbf{H} = \{a_1, \dots, a_K\}$  and  $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_K\}$

with  $\sum_k \alpha_k = \alpha$  and  $\sum_k a_k = 1$ .

- Bernoulli-Multinomial (BMultinomial) model

$$p(\mathbf{f}|\lambda, \mathbf{H}, \boldsymbol{\alpha}) = \prod_j [\lambda \delta(f_j) + (1 - \lambda) \mathcal{M}ult(f_j|\mathbf{H}, \boldsymbol{\alpha})]$$

# Hierarchical models and hidden variables

- All the mixture models and some of simple models can be modeled via **hidden variables  $z$** .

$$p(f) = \sum_{k=1}^K \alpha_k p_k(f) \longrightarrow \begin{cases} p(f|z=k) = p_k(f), \\ P(z=k) = \alpha_k, \quad \sum_k \alpha_k = 1 \end{cases}$$

- Example 1: MoG model:  $p_k(f) = \mathcal{N}(f|m_k, v_k)$   
2 Gaussians:  $p_0 = \mathcal{N}(0, v_0), p_1 = \mathcal{N}(0, v_1), \alpha_0 = \lambda, \alpha_1 = 1 - \lambda$

$$p(f_j|\lambda, v_1, v_0) = \lambda \mathcal{N}(f_j|0, v_1) + (1 - \lambda) \mathcal{N}(f_j|0, v_0)$$

$$\begin{cases} p(f_j|z_j=0, v_0) = \mathcal{N}(f_j|0, v_0), \\ p(f_j|z_j=1, v_1) = \mathcal{N}(f_j|0, v_1), \end{cases} \text{ and } \begin{cases} P(z_j=0) = \lambda, \\ P(z_j=1) = 1 - \lambda \end{cases}$$

$$\begin{cases} p(\mathbf{f}|z) = \prod_j p(f_j|z_j) = \prod_j \mathcal{N}(f_j|0, v_{z_j}) \propto \exp\left[-\frac{1}{2} \sum_j \frac{f_j^2}{v_{z_j}}\right] \\ p(z) = \lambda^{n_1} (1 - \lambda)^{n_0}, \quad n_1 = \sum_j \delta(z_j - 1), \quad n_0 = \sum_j \delta(z_j) \end{cases}$$

# Hierarchical models and hidden variables

- ▶ Example 2: Student-t model

$$St(f|\nu) \propto \exp \left[ -\frac{\nu+1}{2} \log (1 + f^2/\nu) \right]$$

- ▶ Infinite mixture

$$St(f|\nu) \propto= \int_0^\infty \mathcal{N}(f|0, 1/z) \mathcal{G}(z|\alpha, \beta) dz, \quad \text{with } \alpha = \beta = \nu/2$$

$$\begin{cases} p(\mathbf{f}|z) &= \prod_j p(f_j|z_j) = \prod_j \mathcal{N}(f_j|0, 1/z_j) \propto \exp \left[ -\frac{1}{2} \sum_j z_j f_j^2 \right] \\ p(z|\alpha, \beta) &= \prod_j \mathcal{G}(z_j|\alpha, \beta) \propto \prod_j z_j^{(\alpha-1)} \exp [-\beta z_j] \\ p(\mathbf{f}, z|\alpha, \beta) &\propto \exp \left[ -\frac{1}{2} \sum_j z_j f_j^2 + (\alpha-1) \ln z_j - \beta z_j \right] \end{cases}$$

# Hierarchical models and hidden variables

- ▶ Example 3: Laplace (Double Exponential) model

$$\mathcal{DE}(f|a) = \frac{a}{2} \exp[-a|f|] = \int_0^\infty \mathcal{N}(f|0, z) \mathcal{E}(z|a^2/2) dz, \quad a > 0$$

$$\begin{cases} p(\mathbf{f}|z) &= \prod_j p(f_j|z_j) = \prod_j \mathcal{N}(f_j|0, z_j) \propto \exp\left[-\frac{1}{2} \sum_j f_j^2/z_j\right] \\ p(z|\frac{a^2}{2}) &= \prod_j \mathcal{E}(z_j|\frac{a^2}{2}) \propto \exp\left[\sum_j \frac{a^2}{2} z_j\right] \\ p(\mathbf{f}, z|\frac{a^2}{2}) &\propto \exp\left[-\frac{1}{2} \sum_j f_j^2/z_j + \frac{a^2}{2} z_j\right] \end{cases}$$

- ▶ With these models we have:

- ▶ Simple priors

$$p(\mathbf{f}, \boldsymbol{\theta}|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}_1) p(\mathbf{f}|\boldsymbol{\theta}_2) p(\boldsymbol{\theta})$$

- ▶ Hierarchical priors

$$p(\mathbf{f}, z, \boldsymbol{\theta}|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}_1) p(\mathbf{f}|z, \boldsymbol{\theta}_2) p(z|\boldsymbol{\theta}_3) p(\boldsymbol{\theta})$$

# Bayesian Computation and Algorithms

- When the expression of  $p(\mathbf{f}, \boldsymbol{\theta} | \mathbf{g})$  or of  $p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g})$  is obtained, we have following options:
- Joint MAP:** (needs optimization algorithms)

$$(\hat{\mathbf{f}}, \hat{\boldsymbol{\theta}}) = \arg \max_{(\mathbf{f}, \boldsymbol{\theta})} \{p(\mathbf{f}, \boldsymbol{\theta} | \mathbf{g})\}$$

- MCMC:** Needs the expressions of the conditionals  $p(\mathbf{f} | \mathbf{z}, \boldsymbol{\theta}, \mathbf{g})$ ,  $p(\mathbf{z} | \mathbf{f}, \boldsymbol{\theta}, \mathbf{g})$ , and  $p(\boldsymbol{\theta} | \mathbf{f}, \mathbf{z}, \mathbf{g})$
- Variational Bayesian Approximation (VBA):** Approximate  $p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g})$  by a separable one

$$q(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g}) = q_1(\mathbf{f}) q_2(\mathbf{z}) q_3(\boldsymbol{\theta})$$

and do any computations with these separable ones.

# Joint MAP

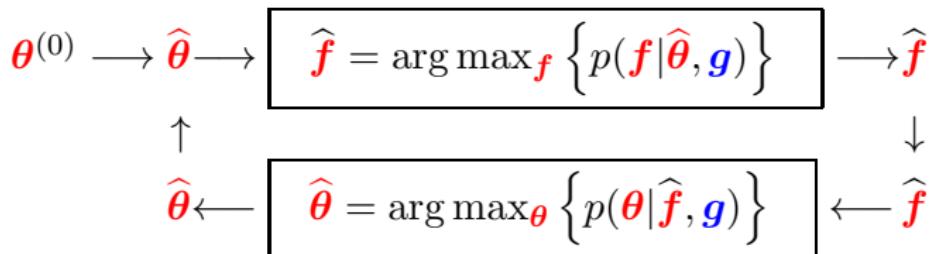
$$p(\mathbf{f}, \boldsymbol{\theta} | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \boldsymbol{\theta}_1) p(\mathbf{f} | \boldsymbol{\theta}_2) p(\boldsymbol{\theta})$$

- ▶ Objective:

$$(\hat{\mathbf{f}}, \hat{\boldsymbol{\theta}}) = \arg \max_{(\mathbf{f}, \boldsymbol{\theta})} \{p(\mathbf{f}, \boldsymbol{\theta} | \mathbf{g})\}$$

- ▶ Alternate optimization:

$$\begin{cases} \hat{\mathbf{f}} = \arg \max_{\mathbf{f}} \{p(\mathbf{f}, \hat{\boldsymbol{\theta}} | \mathbf{g})\} \\ \hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \{p(\hat{\mathbf{f}}, \boldsymbol{\theta} | \mathbf{g})\} \end{cases}$$



- ▶ Uncertainties are not propagated.

# 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 (Gibbs Sampling):

- ▶ Generate samples from the conditionals:

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

- ▶ Wait for convergency
- ▶ Compute empirical statistics (means, modes, variances) from the samples

$$\mathbb{E}\{\mathbf{f}\} \approx \frac{1}{N} \sum_n \mathbf{f}^{(n)}$$

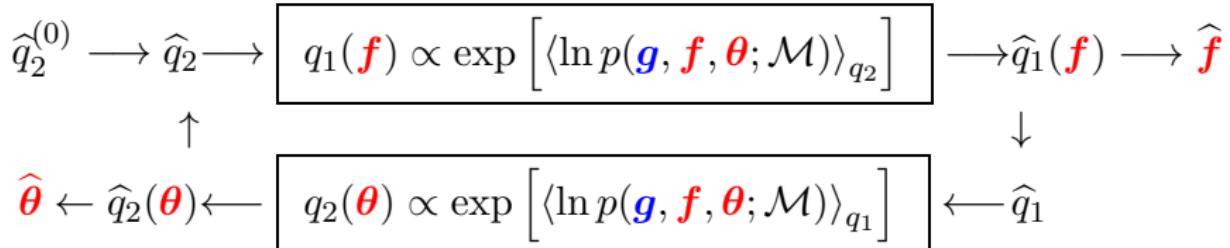
# Variational Bayesian Approximation

- ▶ Approximate  $p(\mathbf{f}, \boldsymbol{\theta} | \mathbf{g})$  by  $q(\mathbf{f}, \boldsymbol{\theta} | \mathbf{g}) = q_1(\mathbf{f} | \mathbf{g}) q_2(\boldsymbol{\theta} | \mathbf{g})$  and then continue computations.
- ▶ Criterion  $\text{KL}(q(\mathbf{f}, \boldsymbol{\theta} | \mathbf{g}) : p(\mathbf{f}, \boldsymbol{\theta} | \mathbf{g}))$

$$\text{KL}(q : p) = \iint q \ln \frac{q}{p} = \iint q_1 q_2 \ln \frac{q_1 q_2}{p}$$

- ▶ Iterative algorithm  $q_1 \rightarrow q_2 \rightarrow q_1 \rightarrow q_2, \dots$

$$\begin{cases} q_1(\mathbf{f}) \propto \exp \left[ \langle \ln p(\mathbf{g}, \mathbf{f}, \boldsymbol{\theta}; \mathcal{M}) \rangle_{q_2(\boldsymbol{\theta})} \right] \\ q_2(\boldsymbol{\theta}) \propto \exp \left[ \langle \ln p(\mathbf{g}, \mathbf{f}, \boldsymbol{\theta}; \mathcal{M}) \rangle_{q_1(\mathbf{f})} \right] \end{cases}$$

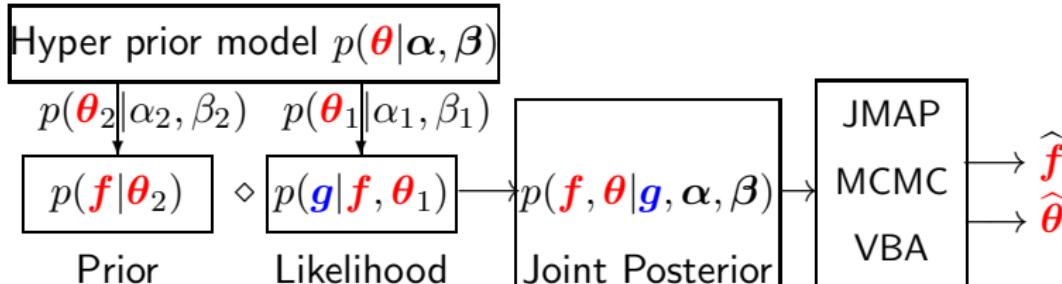


- ▶ Uncertainties are propagated (Message Passing methods)

# Summary of Bayesian approach

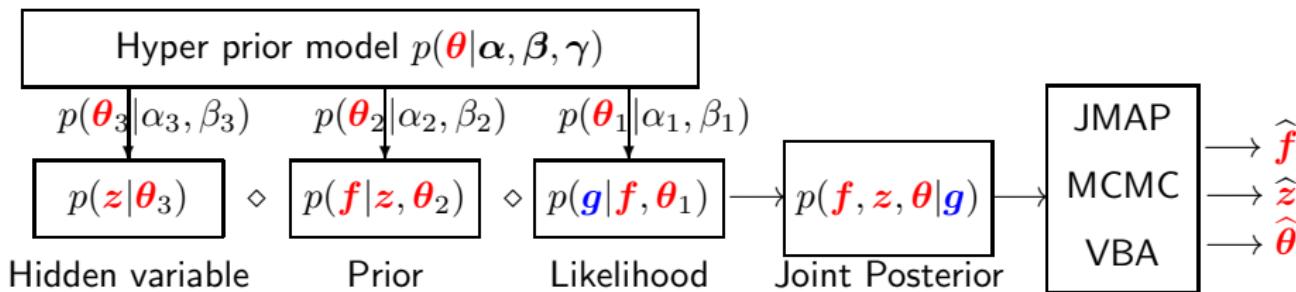
- ▶ Simple priors

$$\downarrow \alpha, \beta$$



- ▶ Hierarchical priors

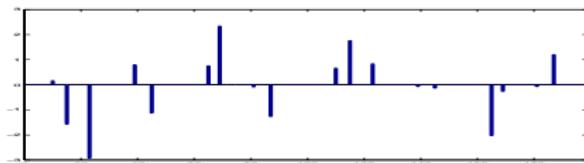
$$\downarrow \alpha, \beta, \gamma$$



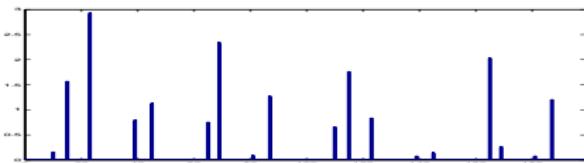
# Advantages of the Bayesian Approach

- ▶ More possibilities to model sparsity
- ▶ More tools to handle hyperparameters
- ▶ More tools to account for uncertainties
- ▶ More possibilities to understand and to control many ad hoc deterministic algorithms
- ▶ Hierarchical models give still more modeling possibilities
  - ▶ Bernouilli-Gaussian: strict sparsity
  - ▶ Bernouilli-Gamma: strict sparsity + positivity
  - ▶ Bernouilli-Multinomial: strict sparsity + discrete values (finite states)
  - ▶ Independent Mixture models: sparsity enforcing
  - ▶ Mixture of multivariate models: group sparsity enforcing
  - ▶ Gauss-Markov-Potts models: sparsity in transform domains

# Examples of Hierarchical models



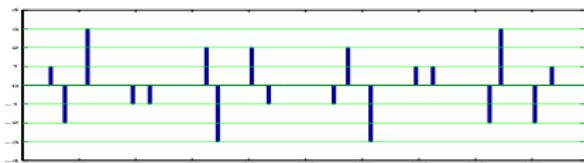
Bernoulli-Gaussian



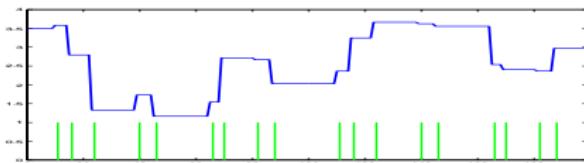
Bernoulli-Gamma



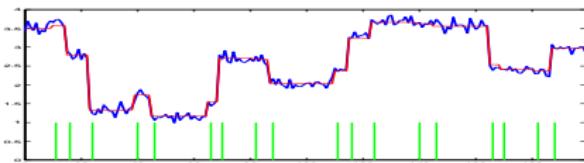
Bernoulli-Binomial



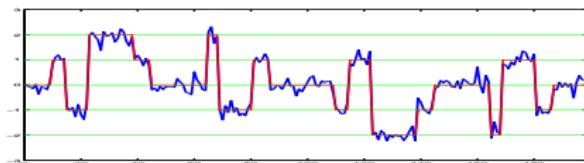
Bernoulli-Multinomial



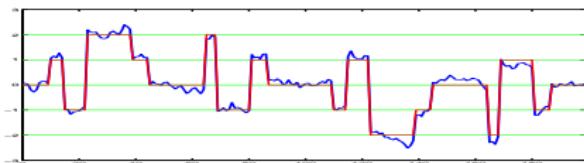
Piecewise constant



Piecewise Gaussian

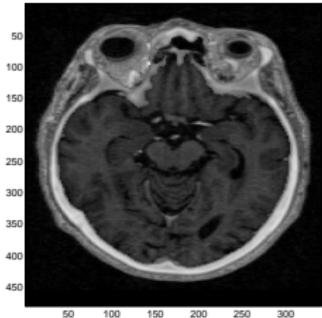


Gauss-Markov-Potts 1



Gauss-Markov-Potts 2

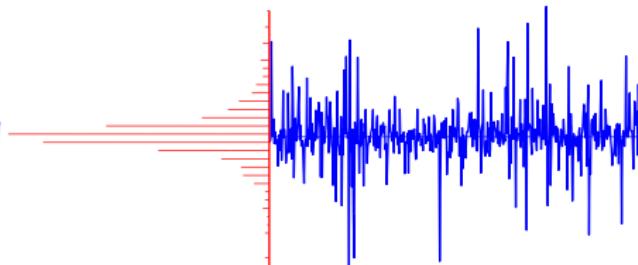
# Which class of images I am looking for?



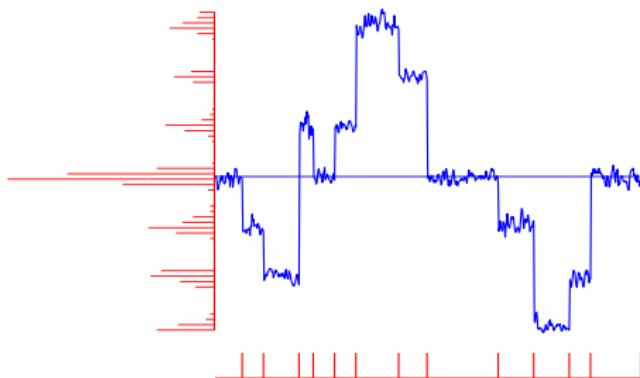
# Which class of signals I am looking for?



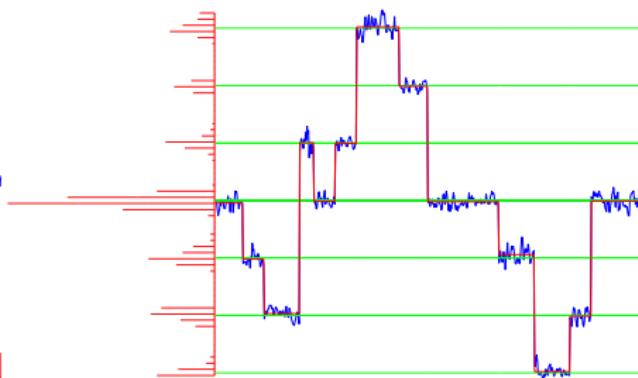
Gauss-Markov



Generalized GM

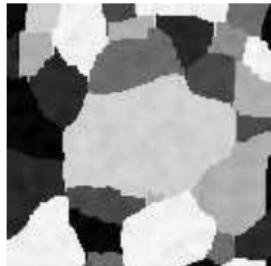
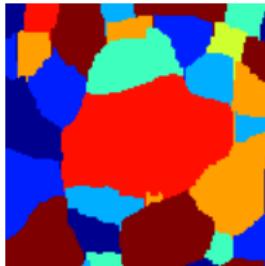


Piecewise Gaussian



Mixture of GM: Gauss-Markov-Potts

# Gauss-Markov-Potts prior models for images

 $f(\mathbf{r})$  $z(\mathbf{r})$ 

$$q(\mathbf{r}) = 1 - \delta(z(\mathbf{r}) - z(\mathbf{r}'))$$

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

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

- ▶ Separable iid hidden variables:  $p(\mathbf{z}) = \prod_{\mathbf{r}} p(z(\mathbf{r}))$
- ▶ Markovian hidden variables:  $p(\mathbf{z})$  Potts-Markov:

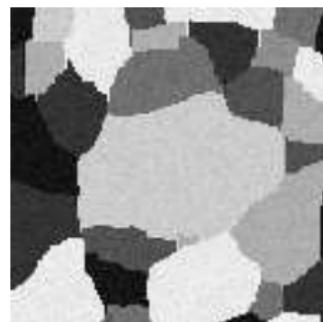
$$p(z(\mathbf{r})|z(\mathbf{r}'), \mathbf{r}' \in \mathcal{V}(\mathbf{r})) \propto \exp \left[ \gamma \sum_{\mathbf{r}' \in \mathcal{V}(\mathbf{r})} \delta(z(\mathbf{r}) - z(\mathbf{r}')) \right]$$
$$p(\mathbf{z}) \propto \exp \left[ \gamma \sum_{\mathbf{r} \in \mathcal{R}} \sum_{\mathbf{r}' \in \mathcal{V}(\mathbf{r})}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

## Four different cases

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

- ▶  $f|z$  Gaussian iid,  $z$  iid :

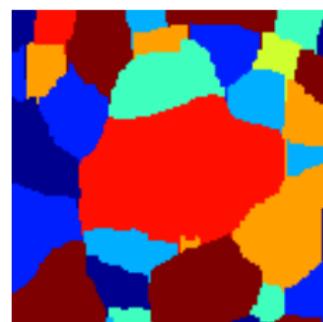
Mixture of Gaussians



$f(\mathbf{r})$

- ▶  $f|z$  Gauss-Markov,  $z$  iid :

Mixture of Gauss-Markov



$z(\mathbf{r})$

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

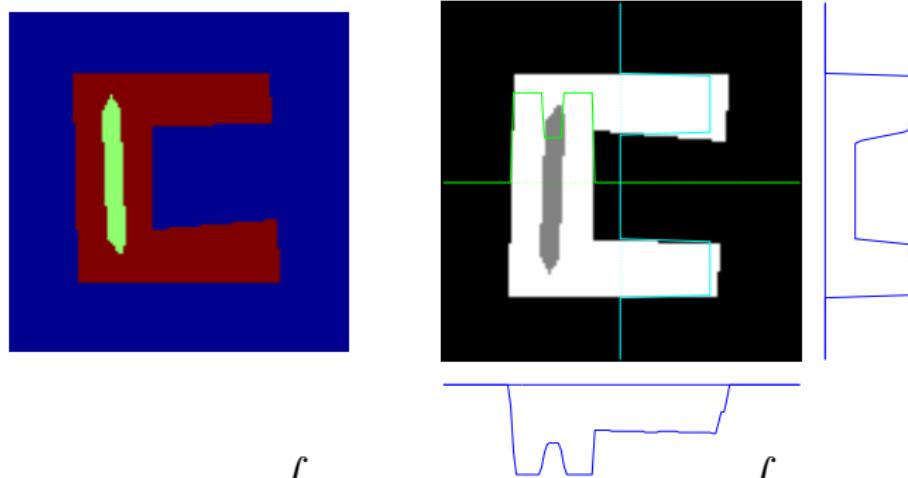
Mixture of Independent Gaussians  
(MIG with Hidden Potts)

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

Mixture of Gauss-Markov  
(MGM with hidden Potts)

# Application of CT in NDT

Reconstruction from only 2 projections

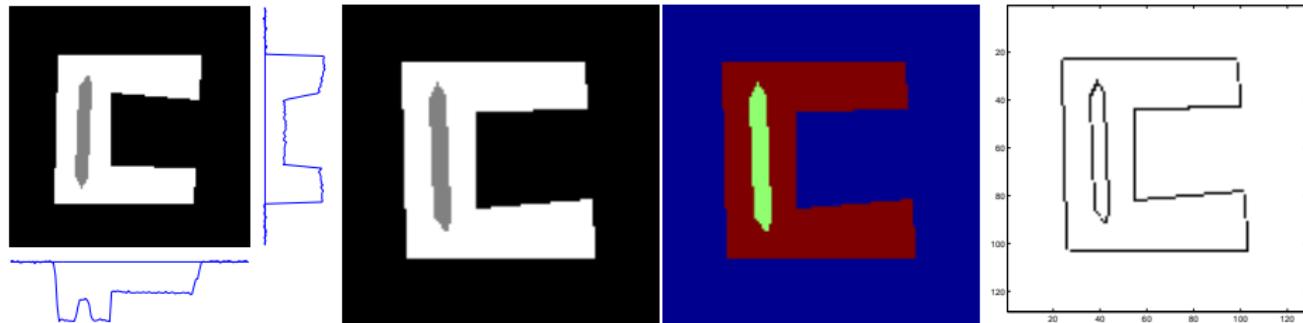


$$g_1(x) = \int f(x, y) \, dy, \quad g_2(y) = \int f(x, y) \, dx$$

- 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



$\mathbf{g} \mathbf{f}$	$\mathbf{f} z$	$z$	$c$
$\mathbf{g} = \mathbf{H}\mathbf{f} + \epsilon$	iid Gaussian	iid	$q(\mathbf{r}) \in \{0, 1\}$
$\mathbf{g} \mathbf{f} \sim \mathcal{N}(\mathbf{H}\mathbf{f}, \sigma_\epsilon^2 \mathbf{I})$	or	or	$1 - \delta(z(\mathbf{r}) - z(\mathbf{r}'))$
Gaussian	Gauss-Markov	Potts	binary

$$p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g}) \propto p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}_1) p(\mathbf{f}|\mathbf{z}, \boldsymbol{\theta}_2) p(\mathbf{z}|\boldsymbol{\theta}_3) p(\boldsymbol{\theta})$$

# Proposed algorithms

$$p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \boldsymbol{\theta}_1) p(\mathbf{f} | \mathbf{z}, \boldsymbol{\theta}_2) p(\mathbf{z} | \boldsymbol{\theta}_3) p(\boldsymbol{\theta})$$

- MCMC based general scheme:

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

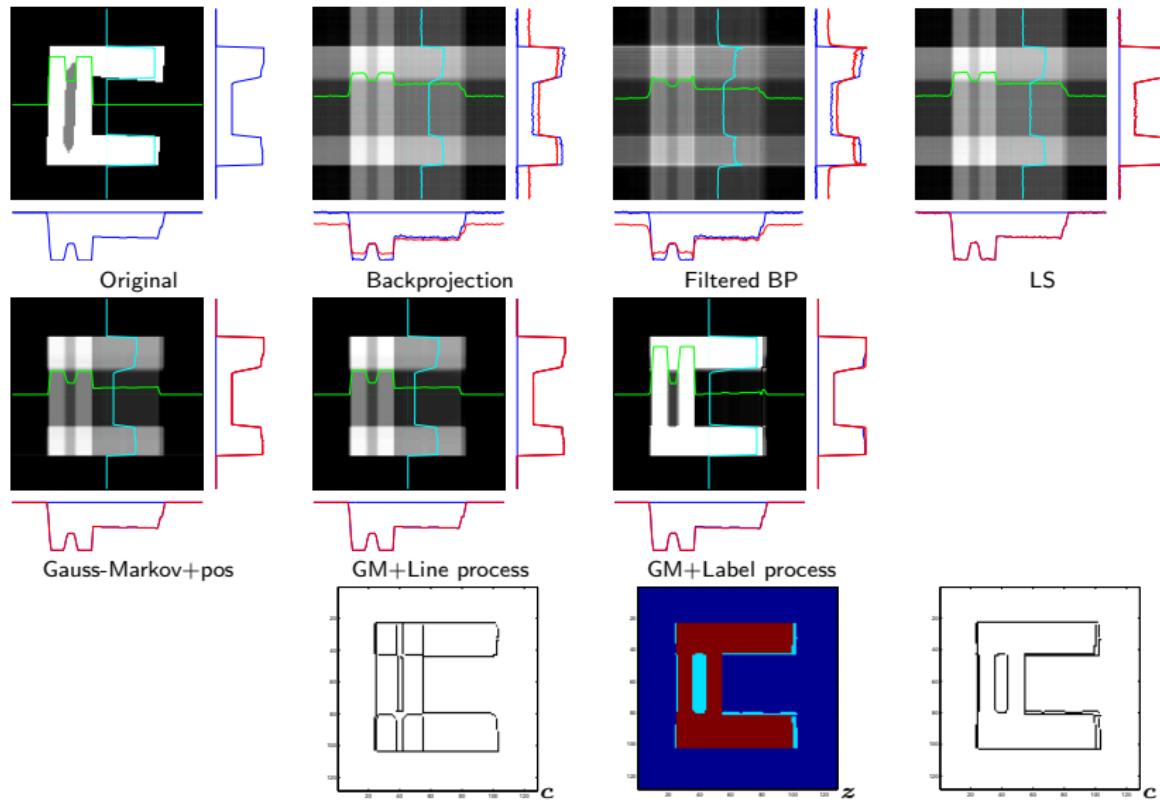
Iterative algorithm:

- ▶ 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.

- Variational Bayesian Approximation

- ▶ Approximate  $p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g})$  by  $q_1(\mathbf{f}) q_2(\mathbf{z}) q_3(\boldsymbol{\theta})$

# Results

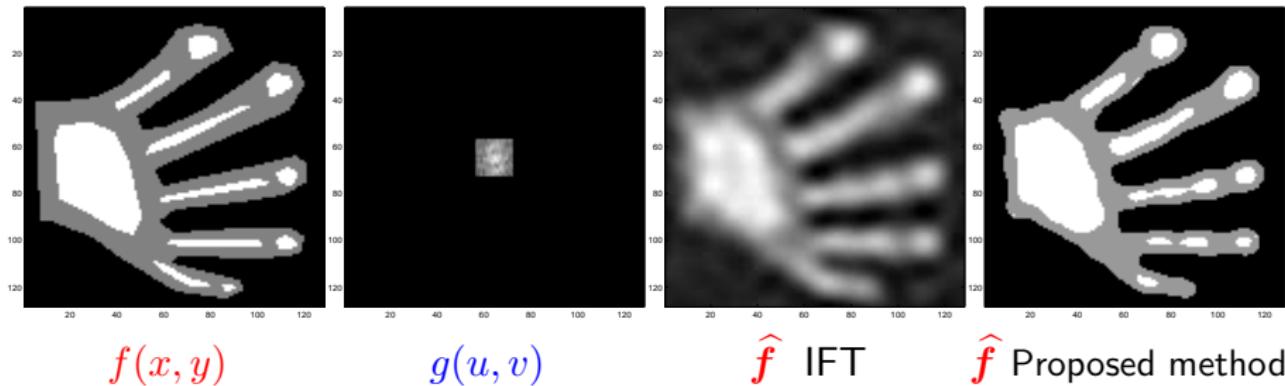


# Application in Microwave imaging

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

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

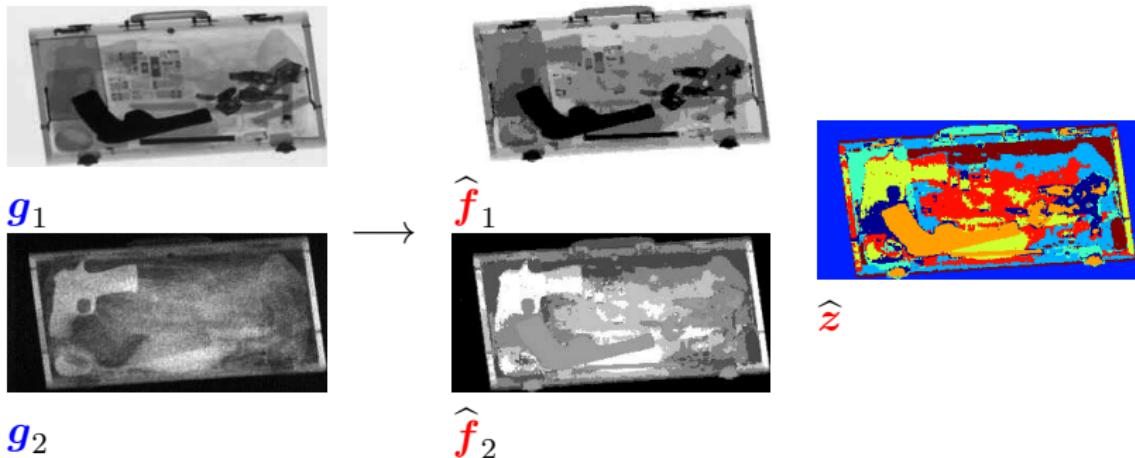
$$\mathbf{g} = \mathbf{H} \mathbf{f} + \boldsymbol{\epsilon}$$



# Images fusion and joint segmentation

(with O. Féron)

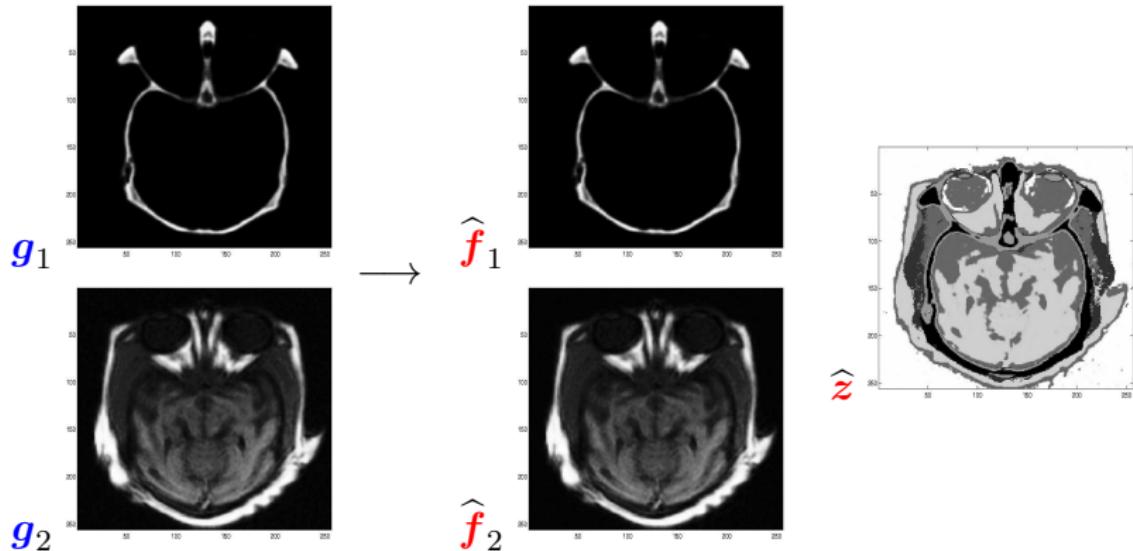
$$\begin{cases} g_i(\mathbf{r}) = f_i(\mathbf{r}) + \epsilon_i(\mathbf{r}) \\ p(f_i(\mathbf{r}) | z(\mathbf{r}) = k) = \mathcal{N}(m_{ik}, \sigma_{ik}^2) \\ p(\underline{\mathbf{f}} | \underline{z}) = \prod_i p(\underline{f}_i | \underline{z}) \end{cases}$$



# Data fusion in medical imaging

(with O. Féron)

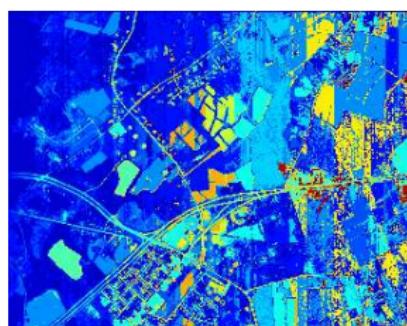
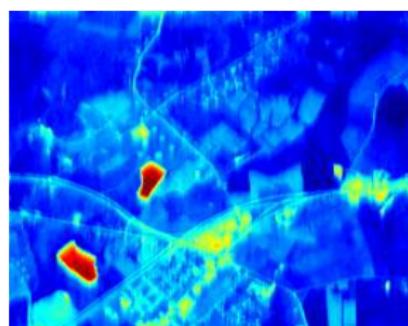
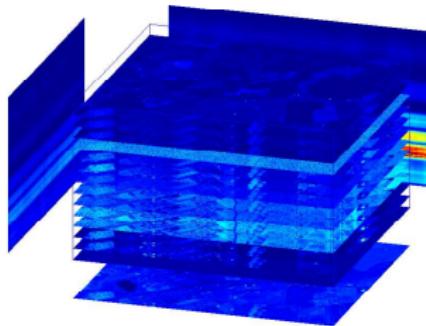
$$\begin{cases} \underline{g_i(\mathbf{r})} = \underline{f_i(\mathbf{r})} + \epsilon_i(\mathbf{r}) \\ p(f_i(\mathbf{r})|z(\mathbf{r}) = k) = \mathcal{N}(m_{ik}, \sigma_{ik}^2) \\ p(\underline{f}|z) = \prod_i p(\underline{f_i}|z) \end{cases}$$



# Joint segmentation of hyper-spectral images

(with N. Bali & A. Mohammadpour)

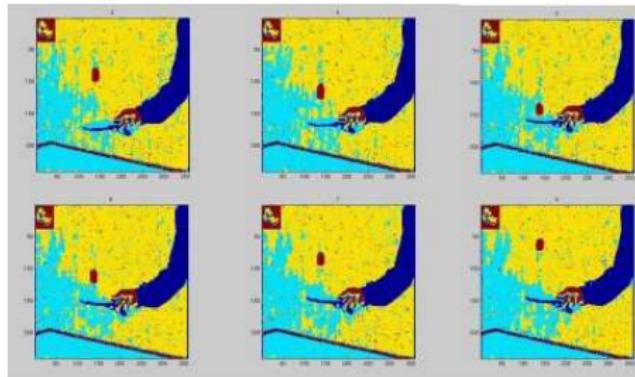
$$\left\{ \begin{array}{l} g_i(\mathbf{r}) = f_i(\mathbf{r}) + \epsilon_i(\mathbf{r}) \\ p(f_i(\mathbf{r})|z(\mathbf{r}) = k) = \mathcal{N}(m_{ik}, \sigma_{ik}^2), \quad k = 1, \dots, K \\ p(\underline{\mathbf{f}}|\underline{\mathbf{z}}) = \prod_i p(\mathbf{f}_i|\mathbf{z}_i) \\ m_{ik} \text{ follow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



# Segmentation of a video sequence of images

(with P. Brault)

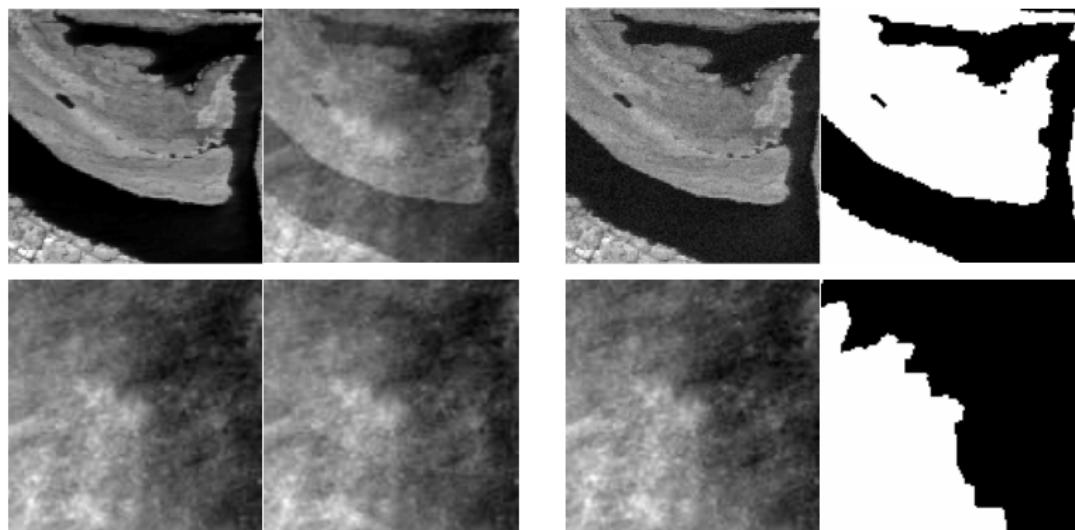
$$\begin{cases} g_i(\mathbf{r}) = f_i(\mathbf{r}) + \epsilon_i(\mathbf{r}) \\ p(f_i(\mathbf{r}) | z_i(\mathbf{r}) = k) = \mathcal{N}(m_{ik}, \sigma_{ik}^2), \quad k = 1, \dots, K \\ p(\underline{\mathbf{f}} | \mathbf{z}) = \prod_i p(\mathbf{f}_i | \mathbf{z}_i) \\ z_i(\mathbf{r}) \text{ follow a Markovian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



# Image separation in Sattelite imaging

(with H. Snoussi & M. Ichir)

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



# Conclusions

- ▶ Sparsity: a great property to use in signal and image processing
- ▶ Origine: Sampling theory and reconstruction, modeling and representation Compressed Sensing, Approximation theory
- ▶ Deterministic Algorithms: Optimization of a two termes criterion, penalty term, regularization term
- ▶ Probabilistic: Bayesian approach
- ▶ Sparsity enforcing priors: Simple heavy tailed and Hierarchical with hidden variables.
- ▶ Gauss-Markov-Potts models for images incorporating hidden regions and contours
- ▶ Main Bayesian computation tools: JMAP, MCMC and VBA
- ▶ Application in different imaging system (X ray CT, Microwaves, PET, ultrasound and microwave imaging)

## Current Projects:

- ▶ Efficient implementation in 2D and 3D cases
- ▶ Comparison between MCMC and VBA methods

# Thanks to:

## Graduated PhD students:

- ▶ Sh. Zhu (2012: SAR Imaging)
- ▶ D. Fall (2012: Emission Positon Tomography, Non Parametric Bayesian)
- ▶ D. Pougaza (2011: Copula and Tomography)
- ▶ H. Ayasso (2010: Optical Tomography, Variational Bayes)
- ▶ S. Fékih-Salem (2009: 3D X ray Tomography)
- ▶ N. Bali (2007: Hyperspectral imaging)
- ▶ O. Féron (2006: Microwave imaging)
- ▶ F. Humblot (2005: Super-resolution)
- ▶ M. Ichir (2005: Image separation in Wavelet domain)
- ▶ P. Brault (2005: Video segmentation using Wavelet domain)
- ▶ H. Snoussi (2003: Sources separation)
- ▶ Ch. Soussen (2000: Geometrical Tomography)
- ▶ G. Montémont (2000: Detectors, Filtering)
- ▶ H. Carfantan (1998: Microwave imaging)
- ▶ S. Gautier (1996: Gamma ray imaging for NDT)
- ▶ M. Nikolova (1994: Piecewise Gaussian models and GNC)
- ▶ D. Prémel (1992: Eddy current imaging)

# Thanks to:

## Freshly graduated (2013) PhD students:

- ▶ C. Cai (Multispectral X ray Tomography)
- ▶ N. Chu (Acoustic sources localization)
- ▶ Th. Boulay (Non Cooperative Radar Target Recognition)
- ▶ R. Prenon (Proteomic and Masse Spectrometry)

## Present PhD students:

- ▶ L. Gharsali (Microwave imaging for Cancer detection)
- ▶ M. Dumitru (Multivariate time series analysis for biological signals)
- ▶ S. AlAli (Diffraction imaging for geophysical applications)

## Present Master students:

- ▶ A. Lupu (Non-circular X ray Tomography)
- ▶ F. Fuc (Multi component signal analysis for biology applications)

## Post-Docs:

- ▶ J. Lapuyade (2011: Dimentionality Reduction and multivariate analysis)
- ▶ S. Su (2006: Color image separation)
- ▶ A. Mohammadpour (2004: HyperSpectral image segmentation)

# Thanks my colleagues and collaborators

- ▶ B. Duchêne & A. Joisel (L2S) (Inverse scattering and Microwave Imaging)
- ▶ N. Gac (L2S) (GPU Implementation)
- ▶ Th. Rodet (L2S) (Computed 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)
- ▶ J. Picheral (SSE, Supélec) (Acoustic sources localization)
- ▶ D. Blacodon (ONERA) (Acoustic sources separation)
- ▶ J. Lagoutte (Thales Air Systems) (Non Cooperative Radar Target Recognition)
- ▶ P. Grangeat (LETI, CEA, Grenoble) (Proteomic and Masse Spectrometry)
- ▶ F. Lévi (CNRS-INSERM, Hopital Paul Brousse) (Biological rythms and Chronotherapy of Cancer)

# References 1

- ▶ A. Mohammad-Djafari, "Bayesian approach with prior models which enforce sparsity in signal and image processing," EURASIP Journal on Advances in Signal Processing, vol. Special issue on Sparse Signal Processing, (2012).
- ▶ A. Mohammad-Djafari (Ed.) Problèmes inverses en imagerie et en vision (Vol. 1 et 2), Hermès-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.
- ▶ 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.
- ▶ H. Ayasso and Ali Mohammad-Djafari Joint NDT Image Restoration and Segmentation using Gauss-Markov-Potts Prior Models and Variational Bayesian Computation, 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.
- ▶ N. Bali and A. Mohammad-Djafari, "Bayesian Approach With Hidden Markov Modeling and Mean Field Approximation for Hyperspectral Data Analysis," IEEE Trans. on Image Processing 17: 2. 217-225 Feb. (2008).
- ▶ H. Snoussi and J. Idier., "Bayesian blind separation of generalized hyperbolic processes in noisy and underdeterminate mixtures," IEEE Trans. on Signal Processing, 2006.

# References 2

- ▶ 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.
- ▶ H. Snoussi and A. Mohammad-Djafari, "Estimation of Structured Gaussian Mixtures: The Inverse EM Algorithm," *IEEE Trans. on Signal Processing* 55: 7. 3185-3191 July (2007).
- ▶ N. Bali and A. Mohammad-Djafari, "A variational Bayesian Algorithm for BSS Problem with Hidden Gauss-Markov Models for the Sources," in: *Independent Component Analysis and Signal Separation (ICA 2007)* Edited by:M.E. Davies, Ch.J. James, S.A. Abdallah, M.D. Plumbley. 137-144 Springer (LNCS 4666) (2007).

## References 3

- ▶ N. Bali and A. Mohammad-Djafari, "Hierarchical Markovian Models for Joint Classification, Segmentation and Data Reduction of Hyperspectral Images" ESANN 2006, September 4-8, Belgium. (2006)
- ▶ M. Ichir and A. Mohammad-Djafari, "Hidden Markov models for wavelet-based blind source separation," IEEE Trans. on Image Processing 15: 7. 1887-1899 July (2005)
- ▶ S. Moussaoui, C. Carteret, D. Brie and A Mohammad-Djafari, "Bayesian analysis of spectral mixture data using Markov Chain Monte Carlo methods sampling," Chemometrics and Intelligent Laboratory Systems 81: 2. 137-148 (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)
- ▶ H. Snoussi and A. Mohammad-Djafari, "Bayesian unsupervised learning for source separation with mixture of Gaussians prior," Journal of VLSI Signal Processing Systems 37: 2/3. 263-279 June/July (2004)
- ▶ F. Su and A. Mohammad-Djafari, "An Hierarchical Markov Random Field Model for Bayesian Blind Image Separation," 27-30 May 2008, Sanya, Hainan, China: International Congress on Image and Signal Processing (CISP 2008).
- ▶ N. Chu, J. Picheral and A. Mohammad-Djafari, "A robust super-resolution approach with sparsity constraint for near-field wideband acoustic imaging," *IEEE International Symposium on Signal Processing and Information Technology* pp 286-289, Bilbao, Spain, Dec14-17,2011

# Questions, Discussions, Open mathematical problems

- ▶ Sparsity representation, low rank matrix decomposition
  - ▶ Sparsity and positivity or other constraints
  - ▶ Group sparsity
  - ▶ Algorithmic and implementation issues for great dimensional applications
- ▶ Optimization of the KL divergence for VBA separable approximation
  - ▶ Convergency of alternate optimization
  - ▶ Other possible algorithms
- ▶ Properties of the obtained approximation
  - ▶ Does the moments of  $q$ 's corresponds to the moments of  $p$ ?
  - ▶ How about any other statistics: entropy, ...
- ▶ Other divergency measures?