Inverse Problems: From Regularization to Bayesian Inference
An Overview on Prior Modeling and Bayesian Computation
Application to Computed Tomography

Ali Mohammad-Djafari

Groupe Problèmes Inverses
Laboratoire des Signaux et Systèmes
UMR 8506 CNRS - SUPELEC - Univ Paris Sud 11
Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette, FRANCE.
djafari@lss.supelec.fr
http://djafari.free.fr
http://www.lss.supelec.fr

Applied Math Colloquium, UCLA, USA, July 17, 2009
Organizer: Stanley Osher
Content

- Image reconstruction in Computed Tomography: An ill posed invers problem
- Two main steps in Bayesian approach: Prior modeling and Bayesian computation
- Prior models for images:
  - Separable Gaussian, GG, ...
  - Gauss-Markov, General one layer Markovian models
  - Hierarchical Markovian models with hidden variables (contours and regions)
  - Gauss-Markov-Potts
- Bayesian computation
  - MCMC
  - Variational and Mean Field approximations (VBA, MFA)
- Application: Computed Tomography in NDT
- Conclusions and Work in Progress
- Questions and Discussion
Computed Tomography: 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 radiographs \( g_\phi(r, z) \)

- **Forward model:**
  Line integrals or Radon Transform

\[
g_\phi(r) = \int_{L(r, \phi)} f(x, y) \, dl + \epsilon_\phi(r)
= \int \int f(x, y) \, \delta(r - x \cos \phi - y \sin \phi) \, dx \, dy + \epsilon_\phi(r)
\]

- **Inverse problem:** Image reconstruction

Given the forward model \( \mathcal{H} \) (Radon Transform) and a set of data \( g_\phi(r), i = 1, \ldots, M \)
find \( f(x, y) \)
2D and 3D Computed Tomography

\[ g_\phi(r_1, r_2) = \int_{L_{r_1, r_2, \phi}} f(x, y, z) \, dl \quad g_\phi(r) = \int_{L_{r, \phi}} f(x, y) \, dl \]

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

Inverse problem: \( g_\phi(r) \) or \( g_\phi(r_1, r_2) \) \( \longrightarrow \) \( f(x, y) \) or \( f(x, y, z) \)
X ray Tomography

\[ g(r, \phi) = -\ln \left( \frac{I}{I_0} \right) = \int_{L_{r, \phi}} f(x, y) \, dl \]

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

\[ f(x, y) \xrightarrow{\text{RT}} g(r, \phi) \]

IRT
Analytical Inversion methods

Radon:

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

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

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

Derivation \( D \):
\[ g(r, \phi) = \frac{\partial g(r, \phi)}{\partial r} \]

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

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

\[ f(x, y) = B \ H \ D \ g(r, \phi) = B \ F_1^{-1} |\Omega| F_1 g(r, \phi) \]

- Backprojection of filtered projections:

\[
\begin{array}{c}
g(r,\phi) \\
\text{FT} \ F_1 \\
\end{array} \xrightarrow{\text{Filter}} \begin{array}{c}
\text{Filter} \\
|\Omega| \\
\end{array} \xrightarrow{\text{IFT}} \begin{array}{c}
g_1(r,\phi) \\
F_1^{-1} \\
\end{array} \xrightarrow{\text{Backprojection}} \begin{array}{c}
\text{Backprojection} \\
B \\
\end{array} \xrightarrow{\text{f}}\]

\[ f(x, y) \]
Limitations: Limited angle or noisy data

- Limited angle or noisy data
- Accounting for detector size
- Other measurement geometries: fan beam, ...
CT as a linear inverse problem

\[ g(s_i) = \int_{L_i} f(r) \, dl_i + \epsilon(s_i) \quad \longrightarrow \quad \text{Discretization} \quad \longrightarrow \quad g = H f + \epsilon \]

- \( g, f \) and \( H \) are huge dimensional
Inversion: Deterministic methods

Data matching

- Observation model
  \[ g_i = h_i(f) + \epsilon_i, \quad i = 1, \ldots, M \rightarrow g = H(f) + \epsilon \]

- Misatch between data and output of the model \( \Delta(g, H(f)) \)
  \[ \hat{f} = \arg \min_f \{ \Delta(g, H(f)) \} \]

- Examples:
  - LS
    \[ \Delta(g, H(f)) = \| g - H(f) \|_2^2 = \sum_i |g_i - h_i(f)|^2 \]
  - \( L_p \)
    \[ \Delta(g, H(f)) = \| g - H(f) \|_p^p = \sum_i |g_i - h_i(f)|^p, \quad 1 < p < 2 \]
  - KL
    \[ \Delta(g, H(f)) = \sum_i g_i \ln \left( \frac{g_i}{h_i(f)} \right) \]

- In general, does not give satisfactory results for inverse problems.
Inversion : Regularization theory

Inverse problems = Ill posed problems

→ Need for prior information

Functional space (Tikhonov) :

\[ g = \mathcal{H}(f) + \epsilon \rightarrow J(f) = \|g - \mathcal{H}(f)\|^2 + \lambda \|Df\|^2 \]

Finite dimensional space (Philips & Towney) :

\[ g = \mathcal{H}(f) + \epsilon \]

- Minimum norme LS (MNLS) :
  \[ J(f) = \|g - \mathcal{H}(f)\|^2 + \lambda \|f\|^2 \]

- Classical regularization :
  \[ J(f) = \|g - \mathcal{H}(f)\|^2 + \lambda \|Df\|^2 \]

- More general regularization :
  \[ J(f) = Q(g - \mathcal{H}(f)) + \lambda \Omega(Df) \]
  or
  \[ J(f) = \Delta_1(g, \mathcal{H}(f)) + \lambda \Delta_2(f, f_\infty) \]

Limitations :

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

\[ \mathcal{M} : \quad g = Hf + \epsilon \]

- Observation model \( \mathcal{M} \) + Hypothesis on the noise \( \epsilon \)
  \[ p(g|f; \mathcal{M}) = p_\epsilon(g - Hf) \]

- A priori information
  \[ p(f|\mathcal{M}) \]

- Bayes:
  \[ p(f|g; \mathcal{M}) = \frac{p(g|f; \mathcal{M}) p(f|\mathcal{M})}{p(g|\mathcal{M})} \]

Link with regularization:

Maximum A Posteriori (MAP):

\[ \hat{f} = \arg \max_f \{ p(f|g) \} = \arg \max_f \{ p(g|f) p(f) \} \]

\[ = \arg \min_f \{ -\ln p(g|f) - \ln p(f) \} \]

with \( Q(g, Hf) = -\ln p(g|f) \) and \( \lambda \Omega(f) = -\ln p(f) \)

But, Bayesian inference is not only limited to MAP
Case of linear models and Gaussian priors

\[ g = Hf + \epsilon \]

- Hypothesis on the noise: \( \epsilon \sim \mathcal{N}(0, \sigma^2_\epsilon I) \)
  \[ p(g|f) \propto \exp \left\{ -\frac{1}{2\sigma^2_\epsilon} \| g - Hf \|^2 \right\} \]

- Hypothesis on \( f \): \( f \sim \mathcal{N}(0, \sigma^2_f (D^t D)^{-1}) \)
  \[ p(f) \propto \exp \left\{ -\frac{1}{2\sigma^2_f} \| Df \|^2 \right\} \]

- A posteriori:
  \[ p(f|g) \propto \exp \left\{ -\frac{1}{2\sigma^2_\epsilon} \| g - Hf \|^2 \frac{1}{2\sigma^2_f} \| Df \|^2 \right\} \]

- MAP:
  \[ \hat{f} = \arg \max_f \{ p(f|g) \} = \arg \min_f \{ J(f) \} \]
  \[ J(f) = \| g - Hf \|^2 + \lambda \| Df \|^2, \quad \lambda = \frac{\sigma^2_\epsilon}{\sigma^2_f} \]

- Advantage: characterization of the solution
  \[ f|g \sim \mathcal{N}(\hat{f}, \hat{P}) \quad \text{with} \quad \hat{f} = \hat{P} H^t g, \quad \hat{P} = (H^t H + \lambda D^t D)^{-1} \]
MAP estimation with other priors:

\[ \hat{f} = \arg \min_f \{ J(f) \} \quad \text{avec} \quad J(f) = \| g - H f \|^2 + \lambda \Omega(f) \]

Separable priors:

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

Markovian models:

\[ p(f_j | f) \propto \exp\left\{ -\alpha \sum_{i \in N_j} \phi(f_j, f_i) \right\} \rightarrow \Omega(f) = \alpha \sum_j \sum_{i \in N_j} \phi(f_j, f_i), \]
MAP estimation with markovien priors

\[ \hat{f} = \arg \min_f \{ J(f) \} \quad \text{with} \quad J(f) = \| g - H f \|^2 + \lambda \Omega(f) \]

\[ \Omega(f) = \sum_j \phi(f_j - f_{j-1}) \]

with \( \phi(t) : \)

Convex functions :

\[ |t|^\alpha, \sqrt{1 + t^2} - 1, \log(\cosh(t)), \quad \left\{ \begin{array}{ll} t^2 & |t| \leq T \\ 2T|t| - T^2 & |t| > T \end{array} \right. \]

or Non convex functions :

\[ \log(1 + t^2), \quad \frac{t^2}{1 + t^2}, \quad \arctan(t^2), \quad \left\{ \begin{array}{ll} t^2 & |t| \leq T \\ T^2 & |t| > T \end{array} \right. \]
MAP estimation with different prior models

\( g = Hf + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma^2_\epsilon I) \)
\( f = Cf + z, z \sim \mathcal{N}(0, \sigma^2_f I) \)
\( f = (I - C)^{-1}z = Dz \)

\( p(g|f) = \mathcal{N}(Hf, \sigma^2_\epsilon I) \)
\( p(f) = \mathcal{N}(0, \sigma^2_f (D^t D)^{-1}) \)

\( p(f|g) = \mathcal{N}(\hat{f}, \hat{P}_f) \)
\( \hat{f} = \arg \min_f \{ J(f) \} \)
\( J(f) = \| g - Hf \|^2 + \lambda \| Df \|^2 \)

\( \hat{f} = \hat{P} H^t g \)
\( \hat{P}_f = (H^t H + \lambda D^t D)^{-1} \)

\( g = Hf + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma^2_\epsilon I) \)
\( f = Wz, z \sim \mathcal{N}(0, \sigma^2_z I) \)
\( g = HWz + \epsilon \)

\( p(g|z) = \mathcal{N}(HWz, \sigma^2_\epsilon I) \)
\( p(z) = \mathcal{N}(0, \sigma^2_z I) \)

\( p(z|g) = \mathcal{N}(\hat{z}, \hat{P}_z) \)
\( \hat{z} = \arg \min_z \{ J(z) \} \)
\( J(z) = \| g - HWz \|^2 + \lambda \| z \|^2 \)

\( \hat{z} = \hat{P}_z H^t W^t g \)
\( \hat{P}_z = (H^t W^t H + \lambda I)^{-1} \)

\( \hat{f} = W \hat{z} \)
\( \hat{P}_f = W \hat{P}_z W^t \)
**MAP estimation with different prior models**

\[
\begin{align*}
\{ & \quad g = H f + \epsilon \\
& \quad f \sim \mathcal{N}(0, \sigma_f^2(D^t D)^{-1}) \} = \left\{ \begin{array}{l}
\quad g = H f + \epsilon \\
\quad f = C f + z \quad \text{with} \quad z \sim \mathcal{N}(0, \sigma_f^2 I) \\
\quad D f = z \quad \text{with} \quad D = (I - C)
\end{array} \right.
\end{align*}
\]

\[
f \mid g \sim \mathcal{N}(\hat{f}, \hat{P}_f) \quad \text{with} \quad \hat{f} = \hat{P}_f H^t g, \quad \hat{P}_f = (H^t H + \lambda D^t D)^{-1}
\]

\[
J(f) = -\ln p(f \mid g) = \| g - H f \|^2 + \lambda \| D f \|^2
\]

\[
\begin{align*}
\{ & \quad g = H f + \epsilon \\
& \quad f \sim \mathcal{N}(0, \sigma_f^2(W W^t)) \} = \left\{ \begin{array}{l}
\quad g = H f + \epsilon \\
\quad f = W z \quad \text{with} \quad z \sim \mathcal{N}(0, \sigma_f^2 I)
\end{array} \right.
\end{align*}
\]

\[
z \mid g \sim \mathcal{N}(\hat{z}, \hat{P}_z) \quad \text{with} \quad \hat{z} = \hat{P}_z W^t H^t g, \quad \hat{P}_z = (W^t H^t H W + \lambda I)^{-1}
\]

\[
J(z) = -\ln p(z \mid g) = \| g - HW z \|^2 + \lambda \| z \|^2 \quad \rightarrow \quad \hat{f} = W \hat{z}
\]

\[z \text{ decomposition coefficients}\]
MAP estimation and Compressed Sensing

\[ \begin{cases} g = Hf + \epsilon \\ f = Wz \end{cases} \]

- \( W \) a code book matrix, \( z \) coefficients
- Gaussian:

\[ p(z) = \mathcal{N}(0, \sigma_z^2 I) \propto \exp \left\{ -\frac{1}{2\sigma_z^2} \sum_j |z_j|^2 \right\} \]

\[ J(z) = -\ln p(z|g) = \|g - HWz\|^2 + \lambda \sum_j |z_j|^2 \]

- Generalized Gaussian (sparsity, \( \beta = 1 \)):

\[ p(z) \propto \exp \left\{ -\lambda \sum_j |z_j|^\beta \right\} \]

\[ J(z) = -\ln p(z|g) = \|g - HWz\|^2 + \lambda \sum_j |z_j|^\beta \]

\[ z = \arg \min_z \{ J(z) \} \longrightarrow \hat{f} = W\hat{z} \]
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

\[ M : \quad g = H f + \epsilon \]

- Forward & errors model : \[ \rightarrow p(g|f, \theta_1; M) \]
- Prior models \[ \rightarrow p(f|\theta_2; M) \]
- Hyperparameters \( \theta = (\theta_1, \theta_2) \) \[ \rightarrow p(\theta|M) \]
- Bayes : \[ \rightarrow p(f, \theta|g; M) = \frac{p(g|f, \theta; M)p(f|\theta; M)p(\theta|M)}{p(g|M)} \]
- Joint MAP : \( (\hat{f}, \hat{\theta}) = \arg \max_{(f, \theta)} \{ p(f, \theta|g; M) \} \)
- Marginalization : \[ \begin{align*}
    p(f|g; M) &= \int p(f, \theta|g; M) \, df \\
    p(\theta|g; M) &= \int p(f, \theta|g; M) \, d\theta 
\end{align*} \]
- Posterior means : \[ \begin{align*}
    \hat{f} &= \int f \, p(f, \theta|g; M) \, df \, d\theta \\
    \hat{\theta} &= \int \theta \, p(f, \theta|g; M) \, df \, d\theta 
\end{align*} \]
- Evidence of the model :

\[ p(g|M) = \int \int p(g|f, \theta; M)p(f|\theta; M)p(\theta|M) \, df \, d\theta \]
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 image I am looking for?

Gaussian
\[ p(f_j | f_{j-1}) \propto \exp \left\{ -\alpha |f_j - f_{j-1}|^2 \right\} \]

Generalized Gaussian
\[ p(f_j | f_{j-1}) \propto \exp \left\{ -\alpha |f_j - f_{j-1}|^p \right\} \]

Piecewise Gaussian
\[ p(f_j | q_j, f_{j-1}) = \mathcal{N} \left( (1 - q_j)f_{j-1}, \sigma_f^2 \right) \]

Mixture of GM
\[ p(f_j | z_j = k) = \mathcal{N} \left( m_k, \sigma_k^2 \right) \]
Gauss-Markov-Potts prior models for images

"In NDT applications of CT, the objects are, in general, composed of a finite number of materials, and the voxels corresponding to each materials are grouped in compact regions"

How to model this prior information?

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

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

\[
p(z(r)|z(r'), r' \in \mathcal{V}(r)) \propto \exp \left\{ \gamma \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}
\]
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
- $f|z$ Gauss-Markov, $z$ iid : Mixture of Gauss-Markov
- $f|z$ 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)
Four different cases

Case 1: Mixture of Gaussians

Case 2: Mixture of Gauss-Markov

Case 3: MIG with Hidden Potts

Case 4: MGM with hidden Potts
Four different cases

\[ f(r) \mid z(r) \]
\[ z(r) \]

\[ f(r) \mid z(r') \]
\[ z(r) \mid z(r') \]

\[ f(r) \mid f(r'), z(r), z(r') \]
\[ z(r) \]

\[ q(r, r') = \{0, 1\} \]

\[ f(r) \mid f(r'), z(r), z(r') \]
\[ z(r) \mid z(r') \]

\[ q(r, r') = \{0, 1\} \]
Case 1: \( f \mid z \) Gaussian iid, \( z \) iid

Independent Mixture of Independent Gaussians (IMIG):

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

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

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

Noting \( \mathcal{R}_k = \{ r : z(r) = k \} \), \( \mathcal{R} = \bigcup_k \mathcal{R}_k \),

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

we have:

\[
p(f \mid z) = \prod_{r \in \mathcal{R}} \mathcal{N}(m_z(r), v_z(r))
\]

\[
p(z) = \prod_r \alpha_z(r) = \prod_k \alpha_k^{\sum_{r \in \mathcal{R}} \delta(z(r) - k)} = \prod_k \alpha_k^{n_k}
\]
Case 2: \( f | z \) Gauss-Markov, \( z \) iid

Independent Mixture of Gauss-Markov (IMGM):

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

\[
\mu_z(r) = \frac{1}{|V(r)|} \sum_{r' \in V(r)} \mu^*_z(r')
\]
\[
\mu^*_z(r') = \delta(z(r') - z(r)) f(r') + (1 - \delta(z(r') - z(r)) m_z(r')
\]
\[
= (1 - c(r')) f(r') + c(r') m_z(r')
\]

\[
p(f | z) \propto \prod_r N(\mu_z(r), \nu_z(r)) \quad \propto \prod_k \alpha_k N(m_k 1, \Sigma_k)
\]
\[
p(z) = \prod_r \nu_z(r) = \prod_k \alpha^n_k
\]

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: \( f \mid z \) Gauss-Markov, \( z \) Potts

Gauss-Markov as in Case 2:

\[
p(f(r) \mid z(r), z(r'), f(r'), r' \in \mathcal{V}(r)) = \mathcal{N}(\mu_z(r), \nu_z(r)), \forall r \in \mathcal{R}
\]

\[
\mu_z(r) = \frac{1}{|\mathcal{V}(r)|} \sum_{r' \in \mathcal{V}(r)} \mu_z^*(r')
\]

\[
\mu_z^*(r') = \delta(z(r') - z(r)) f(r') + (1 - \delta(z(r') - z(r))) m_z(r')
\]

\[
p(f \mid z) \propto \prod_r \mathcal{N}(\mu_z(r), \nu_z(r)) \propto \prod_k \alpha_k \mathcal{N}(m_k^1, \Sigma_k)
\]

Potts-Markov as in Case 3:

\[
p(z) \propto \exp \left\{ \gamma \sum_{r \in \mathcal{R}} \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}
\]
Summary of the two proposed models

\( f \mid z \) Gaussian iid

\( z \) Potts-Markov

(MIG with Hidden Potts)

\( f \mid z \) Markov

\( z \) Potts-Markov

(MGM with hidden Potts)
Bayesian Computation

\[ p(f, z, \theta | g) \propto p(g | f, z, \nu) p(f | z, m, v) p(z | \gamma, \alpha) p(\theta) \]

\[ \theta = \{ \nu, (\alpha_k, m_k, \nu_k), k = 1, \cdot, K \} \quad p(\theta) \quad \text{Conjugate priors} \]

- Direct computation and use of \( p(f, z, \theta | g; \mathcal{M}) \) 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
  \[ p(f, z, \theta | g; \mathcal{M}) \quad \text{by} \quad q(f, z, \theta) = q_1(f) q_2(z) q_3(\theta) \]

  - Choice of approximation criterion: \( KL(q : p) \)
  - Choice of appropriate families of probability laws for \( q_1(f) \), \( q_2(z) \) and \( q_3(\theta) \)
MCMC based algorithm

\[ p(f, z, \theta \mid g) \propto p(g \mid f, z, \theta) p(f \mid z, \theta) p(z) p(\theta) \]

General scheme:

\[ \hat{f} \sim p(f \mid \hat{z}, \hat{\theta}, g) \longrightarrow \hat{z} \sim p(z \mid \hat{f}, \hat{\theta}, g) \longrightarrow \hat{\theta} \sim (\theta \mid \hat{f}, \hat{z}, g) \]

- Sample \( f \) from \( p(f \mid \hat{z}, \hat{\theta}, g) \propto p(g \mid f, \theta) p(f \mid \hat{z}, \hat{\theta}) \)
  Needs optimisation of a quadratic criterion.

- Sample \( z \) from \( p(z \mid \hat{f}, \hat{\theta}, g) \propto p(g \mid \hat{f}, \hat{z}, \hat{\theta}) p(z) \)
  Needs sampling of a Potts Markov field.

- Sample \( \theta \) from
  \[ p(\theta \mid \hat{f}, \hat{z}, g) \propto p(g \mid \hat{f}, \sigma^2 \epsilon I) p(f \mid \hat{z}, (m_k, v_k)) p(\theta) \]
  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: $f(x, y) = g_1(x) \cdot g_2(y) \cdot \Omega(x, y)$

$\Omega(x, y)$ is a Copula:

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

\[ g | f \]
\[ g = H f + \epsilon \]
\[ g | f \sim \mathcal{N}(H f, \sigma^2 I) \]
Gaussian

\[ f | z \]
\[ \text{iid Gaussian} \]

\[ z \]
\[ \text{iid} \]

\[ q \]
\[ q(r) \in \{0, 1\} \]
\[ 1 - \delta(z(r) - z(r')) \]
binary

Forward model | Gauss-Markov-Potts Prior Model | Auxiliary

Unsupervised Bayesian estimation:

\[ p(f, z, \theta | g) \propto p(g | f, z, \theta) p(f | z, \theta) p(\theta) \]
Results: 2D case

Original

Backprojection

Filtered BP

LS

Gauss-Markov+pos

GM+Line process

GM+Label process
Some results in 3D case

M. Defrise

FeldKamp

Phantom

Proposed method
Some results in 3D case

FeldKamp

Proposed method
Some results in 3D case

A photograpy of metalique esponge

Experimental setup

Reconstruction by proposed method

Feldkamp

EM 2D

Notre méthode
Application: liquid evaporation in metallic sponge
Conclusions

- Bayesian estimation approach gives more tools than deterministic regularization to handle inverse problems.
- Gauss-Markov-Potts are useful prior models for images incorporating regions and contours.
- Bayesian computation needs either multi-dimensional optimization or integration methods.
- Different optimization and integration tools and approximations exist (Laplace, MCMC, Variational Bayes).

Work in Progress and Perspectives:

- Efficient implementation in 2D and 3D cases using GPU.
- Application to other linear and non-linear inverse problems: X ray CT, Ultrasound and Microwave imaging, PET, SPECT, ...